Juha Ritala

Computational study of quantum dot qubits using Lagrange mesh method and exact diagonalization

School of Science

Thesis submitted for examination for the degree of Master of Science in Technology.

Espoo 25.11.2013

Thesis supervisor:

Prof. Risto Nieminen

Thesis advisor:

D.Sc. (Tech.) Ari Harju

Aalto University
School of Science

AALTO UNIVERSITY
SCHOOL OF SCIENCE

ABSTRACT OF THE
MASTER'S THESIS

Author: Juha Ritala

Title: Computational study of quantum dot qubits using Lagrange mesh method and exact diagonalization

Date: 25.11.2013 Language: English Number of pages:5+88

Department of Applied Physics

Professorship: Engineering Physics Code: Tfy-105

Supervisor: Prof. Risto Nieminen

Advisor: D.Sc. (Tech.) Ari Harju

Quantum computation using quantum circuit model is based on quantum bits and gates, which are quantum analogues to the bits and logical gates in classical computing. The computations are carried out by performing single- and two-qubit quantum gate operations on the input qubits (quantum bits). In this thesis, a set of computational methods to study these gate operations in the case of semiconductor quantum dot based spin qubits is presented. This set consists of three parts. Lagrange mesh method is used to calculate the single-electron states in a quantum dot system. These states are then used in an exact diagonalization calculation to obtain the many-electron ground state, which is then evolved using exact diagonalization based dynamics. The presented set of methods is used to simulate single-qubit gates, and it is found to be successful for this purpose. The Lagrange mesh method is extremely versatile as it can handle an arbitrary quantum dot confinement potential without the need to calculate any integrals. This feature is achieved by approximating the potential matrix element integrals using a Gauss quadrature. The high accuracy of the Lagrange mesh method despite the seemingly crude approximation is investigated, and a reasonable cause for it in the case of low degree polynomial potentials is found. A hypothesis that the Gauss quadrature approximation is extremely accurate for an arbitrary polynomial potential is made. The convergence of the states calculated with the Lagrange mesh method is tested and compared to an alternative method based on localized Gaussian basis functions.

Keywords: quantum computing, qubit, quantum dot,
Lagrange mesh, exact diagonalization

AALTO-YLIOPISTO
PERUSTIETEIDEN KORKEAKOULU

DIPLOMITYÖN
TIIVISTELMÄ

Tekijä: Juha Ritala

Työn nimi: Kvanttipistekubittien laskennallinen tutkimus käyttäen Lagrangen hilan menetelmää ja eksaktia diagonalisointia

Päivämäärä: 25.11.2013 Kieli: Englanti Sivumäärä:5+88

Teknillisen fysiikan laitos

Professuuri: Teknillinen fysiikka Koodi: Tfy-105

Valvoja: Prof. Risto Nieminen

Ohjaaja: TkT Ari Harju

Kvanttipiirimallin mukainen kvanttilaskenta perustuu kvanttibitteihin ja -portteihin, jotka ovat vastine klassisen laskennan biteille ja loogisille porteille. Laskenta suoritetaan tekemällä yhden tai kahden qubitin kvanttiporttioperaatioita syötekubiteille (kvanttibiteille). Tässä diplomityössä esitetään joukko laskennallisia menetelmiä, joilla voidaan tutkia näitä porttioperaatioita, kun kyseessä ovat puolijohdekvanttipisteisiin perustuvat spin-kubitit. Tämä menetelmien joukko koostuu kolmesta osasta. Lagrangen hilan menetelmää käytetään yksielektronitilojen laskemiseen kvanttipistesysteemissä. Näitä tiloja käytetään eksaktissa diagonalisoinnissa, jolla lasketaan usean elektronin perustila. Tätä perustilaa kehitetään eksaktiin diagonalisointiin perustuvalla dynamiikkamenetelmällä. Esitettyjä menetelmiä käytetään simuloimaan yhden kubitin portteja, ja niiden huomataan sopivan hyvin tähän tehtävään. Lagrangen hilan menetelmä on erittäin monipuolinen, koska sillä voidaan käsitellä mielivaltaista kvanttipisteen potentiaalia laskematta yhtään integraalia. Tämä ominaisuus saavutetaan approksimoimalla potentiaalimatriisialkioiden integraalit Gaussin kvadratuurilla. Lagrangen hilan menetelmän korkea tarkkuus tästä näennäisesti karkeasta approksimaatiosta huolimatta on tutkinnan kohteena, ja sille löytyy järkevä syy alhaisen asteen polynomipotentiaalien tapauksessa. Hypoteesi, että Gaussin kvadratuuri -approximaatio on äärimmäisen tarkka mielivaltaiselle polynomipotentiaalille esitetään. Lagrangen hilan menetelmällä laskettujen tilojen suppenemista testataan ja verrataan vaihtoehtoiseen menetelmään, joka perustuu lokalisoituihin Gaussisiin kantafunktioihin.

Avainsanat: kvanttilaskenta, kubitti, kvanttipiste,
Lagrange mesh, eksakti diagonalisointi

Preface

All the research that is included in this thesis was conducted in the Quantum Many-Body Physics (QMP) group of the Department of Applied Physics at Aalto University. I used ready-made program codes that have all been written in the QMP group, and modified and optimized them to suit the needs of this work. I wish to thank my instructor Ari Harju for giving me lots of ideas for this thesis, but letting me to decide which ones to implement. A large part of the work was done in collaboration with my colleague Tuukka Hiltunen, whom I wish to thank for our helpful discussions and for his immense support at the writing phase. Finally, I wish to thank Verna for her unlimited patience during the long working hours and for the vital mental support at those moments when finishing this thesis seemed to be a hopeless task.

Otaniemi, 25.11.2013

Juha Ritala

Contents

Chapter 1

Introduction

The concept of quantum computing was introduced in the beginning of the 1980s. In 1982, Feynman proposed that computers based on the principles of quantum mechanics could be used to avoid the difficulties in simulating quantum systems on classical computers [1]. The difficulties which Feynman refers to lead to the fact that simulating a quantum mechanical system on a conventional computer is an extremely heavy task, and only certain systems can be simulated efficiently. These systems are usually either one or two-dimensional and have up to a few particles. Therefore, one of the goals of quantum computation is to develop tools to efficiently study quantum mechanical effects [2]. Another goal is to solve problems outside quantum mechanics that are too complex for classical computers. Two well-known examples are the Shor's algorithm for factoring large numbers [3] and the Grover's algorithm for searching an unsorted database [4].

Several different models for quantum computers have been suggested since the proposition of Feynman. The model that is closest to a conventional classical computer is the quantum circuit, which has quantum analogues for bits, gates and wires of a classical computer. Quantum circuit has obtained most of the research interest so far and a lot of progress has been made to realize it. The quantum circuit computing model is independent of the physical realization of the quantum bits, gates and wires that are needed for the computer. In this thesis, quantum bits, also known as qubits, that are based on the singlet and triplet spin states of two electrons are studied by computational means. The physical objective of the thesis is simulating the ways to control these qubits, i.e. studying the implementation of the quantum gates that act on single qubits. The simulations are based on the recent developments in the area. However, the main topic is not the simulation of the quantum gates, but the analysis of the method that is an important part of the computational toolkit, the Lagrange mesh method.

The electrons that are used as spin qubits are trapped in so called quantum dots which can be tuned to hold only one electron in them. Quantum dots based on electrostatic confinement of the electrons in the two dimensional electron gas that forms between layers of GaAs and AlGaAs were chosen as the model devices. These devices are relatively easy to fabricate using the present technology and they can be readily tuned by adjusting the voltages at the electrostatic gates. From the compu-

tational point of view, the GaAs-AlGaAs quantum dots are a good choice because they are basically two dimensional devices. As was already noted, exact simulation of a quantum mechanical system using a classical computer is computationally intensive, and only some simple systems can be studied. Luckily, there are efficient methods for simulating two dimensional systems with up to a few electrons using the constantly growing computational resources.

The simulations that are done in this thesis consist of three separate steps. First, the single-electron states in a quantum dot system are solved using the Lagrange mesh method. Then, a many-body basis is constructed from the single-electron states and that basis together with the interaction matrix elements between the single-electron states are used in an exact diagonalization calculation to obtain the ground state of a two-electron system. Finally, the ground state is evolved by running exact diagonalization based dynamics.

This thesis is divided into four main chapters. In Chapter 2, some theoretical background about quantum dots, quantum computing and singlet-triplet qubits is given. The computational tools that are used in the simulations are introduced in Chapter 3. Chapter 4 contains mathematical analysis about the Gauss quadrature approximation on which the Lagrange mesh method relies. All the computational results are presented in Chapter 5, including evaluation of the used methods and the simulations of the qubit control.

Chapter 2

Theoretical background

2.1 Semiconductor quantum dots

Quantum dots are nanoelectronic devices that contain a controllable number of electrons in them. The number of electrons can be accurately controlled because the quantum dot is separated from the surrounding electron reservoir by a tunnel barrier and the available electron states in the dot have discrete energies. This feature is achieved by spatial confinement of the electrons in all three dimensions, which causes quantization of the electron energy states. The confining potential of the quantum dot can be tuned in such a way that a desired number of electron states is below the Fermi level of the reservoir, which results in occupation of those states by electrons that tunnel from the reservoir.

The quantization of the electron states in quantum dots happens due to two separate effects, Coulomb blockade and the quantization of the single-electron states. The Coulomb blockade refers to the blocking mechanism of electrons due to the quantization of the Coulomb interaction energy between electrons. The quantization of the Coulomb interaction is caused by the inherent discrete nature of the electron charge. Therefore, it is essentially a classical effect. The name "quantum dot" refers particularly to the quantization of the single-electron orbitals due to the confinement of the electrons to the length scale of their de Broglie wavelength. This quantum confinement in all spatial dimensions leaves zero classical degrees of freedom, and, in that sense, the device is zero dimensional, i.e. a dot. [5]

The quantum dots have also been called artificial atoms because of the resemblance between atoms and quantum dots, electron confining potential and the quantization of the electron states [6]. The electrical properties of quantum dots differ notably from real atoms though, mainly because of the fact that they are typically much larger than atoms. The separations of the single-electron orbital energies decrease faster than the strength of the Coulomb interaction between the electrons when the size of an atom is increased. Since quantum dots are larger than real atoms, electron-electron interactions are relatively more important in them.

2.1.1 Fabrication

Quantum dots can be fabricated using various different methods, but for the purpose of making spin qubits for quantum computers, quantum dots based on electrostatic confinement of the two dimensional electron gas that forms in semiconductor heterostructures have been proven to be successful. In this thesis, semiconductor quantum dots based on GaAs-AlGaAs heterostructure are considered.

Modern semiconductor technology allows growing of gallium arsenide (GaAs) crystals layer by layer. Layers that have some of the gallium atoms replaced with aluminium can be grown on top of the pure GaAs layers because the lattice constants of GaAs and AlAs almost match [7]. Since GaAs has a smaller band gap than AlGaAs, the conduction band bends at their interface to form a narrow potential well [8]. The electron states in the well are quantized because of the spatial confinement, and the separation of the energies is so large that in low temperatures only the ground state is occupied [6]. The electrons in the well can move freely parallel to the interface, because there is still a continuum of available states in that plane, but not perpendicular to it, and thus they form a two dimensional electron gas [9]. The two dimensional electron gas can be further confined in the two available dimensions using metallic electrodes that are deposited on the surface of the semiconductor heterostructure to form quantum dots. A double quantum dot device based on GaAs-AlGaAs heterostructure is shown in Fig. 2.1.

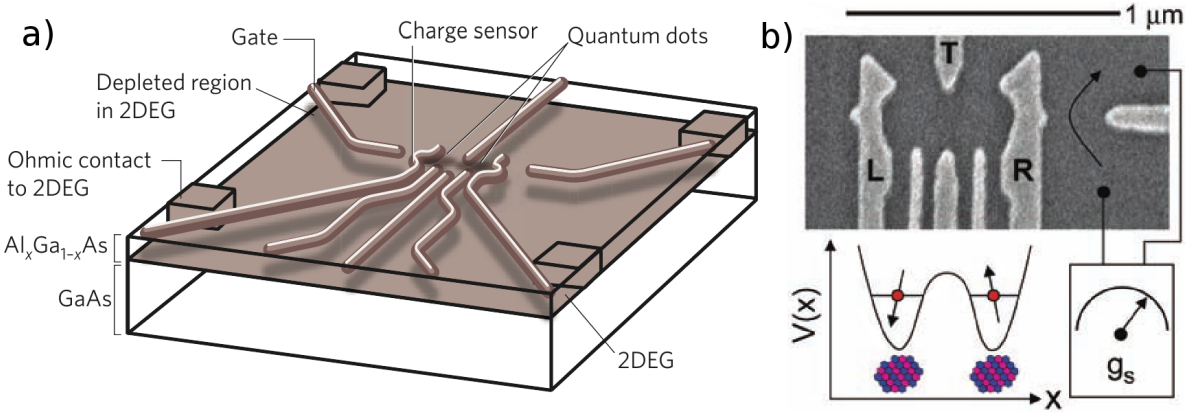

Figure 2.1: a) Schematic picture of a double quantum dot device based on GaAs-AlGaAs heterostructure [10]. b) Scanning electron micrograph of a real double quantum dot device used by Petta et al. [11]. Gates L and R can be used to control the levels of the left and right potential minima respectively. Gate T is used to control the height of the barrier between the two minima. The quantum point contact conductance g_s is primarily sensitive to the charge in the right dot.

2.1.2 Computational model

The electrons that are confined to the quantum well at the interface of the semiconductors interact with the nuclei and the bound electrons of the neighbouring

semiconductor atoms. Therefore, they cannot be treated as if they were free electrons in a vacuum. They can be treated as if they were free, however, by giving them an effective mass that takes into account all the interactions with the surrounding nuclei and core electrons. The permittivity of GaAs must also be taken into account when the Coulomb interaction between the electrons is considered. The effective mass of electrons in GaAs is $m^* \approx 0.067\, m_e$ and the permittivity of GaAs is $\epsilon \approx 12.7\, \epsilon_0$, where m_e is the true mass of the electron and ϵ_0 is the permittivity of the vacuum. An electron in a quantum dot that is placed in a uniform magnetic field can be described using the Hamiltonian

$$H_0 = \frac{(\mathbf{p} + e\mathbf{A}(\mathbf{r}))^2}{2m^*} + V(\mathbf{r}) + V_Z(\mathbf{S}), \tag{2.1}$$

where $\mathbf{A}(\mathbf{r})$ is a vector potential corresponding to the magnetic field and $V_Z(\mathbf{S})$ is the Zeeman energy due to coupling between the electron spin and the magnetic field. The Zeeman term can be written as

$$V_Z(\mathbf{S}) = \frac{g^* \mu_B}{\hbar} \mathbf{B} \cdot \mathbf{S}, \tag{2.2}$$

where g^* is the effective g-factor, which is $g^* = -0.44$ for GaAs, and μ_B is the Bohr magneton. If the uniform magnetic field is chosen to point to the z-direction, the corresponding vector potential can be taken to be $\mathbf{A}(\mathbf{r}) = \frac{B}{2}(y\mathbf{e}_x - x\mathbf{e}_y)$ and the Zeeman term simplifies to $V_Z(S_z) = \frac{g^* \mu_B}{\hbar} B S_z$.

The electrostatic potential that confines the electrons in the plane of the two dimensional electron gas is assumed to be parabolic in the neighbourhood of the potential minimum. The parabolic form of the potential was adopted after Kumar et al. showed that the confinement potential in their device had almost circular symmetry despite the fact that the confining electrode was square-shaped [5, 8]. The parabolic potential can be written as

$$V(\mathbf{r}) = \frac{1}{2} m^* \omega_0^2 r^2, \tag{2.3}$$

where ω_0 is the confinement strength of the potential. The most straightforward way to generalize the parabolic potential approximation for a system of several quantum dots is to use a piecewise parabolic potential,

$$V(\mathbf{r}) = \frac{1}{2} m^* \omega_0^2 \min_{1 \le j \le n} \left\{ |\mathbf{r} - \mathbf{R}_j|^2 \right\}, \tag{2.4}$$

where $\{\mathbf{r}_j\}_{j=1}^n$ are the locations of the parabolic minima. Recently, Nielsen et al. [12] solved the electrostatic potential of an accurately modelled double quantum dot system and they noticed that it indeed was relatively close to the piecewise parabolic potential defined in Eq (2.4). However, the min function causes sharp kinks to the potential at the interfaces of the parabola. These kinks cannot exist in a physical system because the derivatives of the potential are discontinuous at the kinks. The kinks also hinder the convergence of the results, but they can be easily smoothed using a method that is discussed in the results chapter.

Effective atomic units are used in all calculations to simplify the notation and, consequently, the computations by removing most of the physical constants. In effective atomic units, the constants m^*, e, $\hbar$ and $4\pi\epsilon$ are all set to 1. Consequently, energy is measured in units of $\mathrm{Ha}^* \approx 11.30$ meV (effective Hartree) and length in units of $a_0^* = 10.03$ nm (effective Bohr radius). Furthermore, the unit conversion from Teslas to atomic units is given by $1\,\mathrm{T} \approx 0.1529\,\mathrm{a.u.}^*$. By transferring to the atomic units and writing the quantum mechanical momentum operator as $\mathbf{p} = -i\hbar\nabla$, the Hamiltonian can be simplified to

$$H_0 = -\frac{1}{2}\nabla^2 + V(\mathbf{r}) + \frac{1}{8}\omega_c^2 r^2 - i\frac{\omega_c}{2}(y\partial_x - x\partial_y) + V_Z(S_z), \tag{2.5}$$

where $\omega_c = eB/m^* = B$ a.u.*. The Zeeman term is now $V_Z(S_z) = g^*\mu_B B S_z$, where $\mu_B = \frac{1}{2}$ in effective atomic units.

The Hamiltonian for a many-electron system consists of a sum of single-electron Hamiltonians and a term that describes the Coulomb interaction between the electrons. The electron-electron interaction term can be written as

$$U = \sum_{j<k} \frac{e^2}{4\pi\epsilon|\mathbf{r}_j - \mathbf{r}_k|}. \tag{2.6}$$

The complete many-electron Hamiltonian can thus be written as

$$H = \sum_j H_0^{(j)} + \sum_{j<k} \frac{1}{|\mathbf{r}_j - \mathbf{r}_k|}, \tag{2.7}$$

where effective atomic units have been used for the interaction term.

2.2 Quantum computation

Qubits, or quantum bits, are the quantum equivalent of bits in classical information. The qubits have two basis states which are generally denoted as $|0\rangle$ and $|1\rangle$ corresponding to the possible values of a classical bit, 0 and 1. Unlike a classical bit, however, the state of a quantum bit can be a linear combination, or superposition, of the two basis states. If $|\psi\rangle$ is the state of the qubit, it can be written as

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle, \tag{2.8}$$

where α and β are complex numbers that satisfy the normalization condition $|\alpha|^2 + |\beta|^2 = 1$. Because of the normalization condition, the qubit state can be rewritten as

$$|\psi\rangle = e^{i\gamma}\left(\cos\frac{\theta}{2}|0\rangle + e^{i\phi}\sin\frac{\theta}{2}|1\rangle\right). \tag{2.9}$$

Since the overall complex phase of the state, $e^{i\gamma}$, has no physical meaning, the state of the qubit is defined by the real numbers θ and ϕ. These numbers define a point on a unit sphere, which is known as the Bloch sphere in this context, and thus the state

of the qubit can be conveniently represented as that point [2]. The Bloch sphere for a qubit with singlet and triplet states of a two-electron system as basis states is shown in Fig. 2.2.

In classical computers, the computations are carried out by applying logic gates to the bits of the input data. A logic gate performs a logical operation, which depends on the type of the gate, on one or several input bits transforming them to corresponding output bits. In this respect, quantum circuit computers operate similarly as the classical ones, by performing gate operations on the input qubits. Since the state of a qubit is a superposition of two quantum mechanical states, the quantum gate operations are substantially more complex than the classical ones though. Fortunately, only single qubit operations and a two-qubit gate are needed for a functioning quantum computer because of the important universality result: any multiple qubit logic gate may be composed from two-qubit CNOT gate and single qubit gates [2].

As the single qubit gates act linearly on the state of the qubit, they can be represented in a matrix form. In this matrix representation, the state of a qubit $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ can be written as a vector

$$|\psi\rangle \rightarrow \begin{bmatrix} \alpha \\ \beta \end{bmatrix},$$

and the single qubit gates are 2×2 square matrices. The only restriction for a single qubit gate is that the corresponding matrix must be unitary to conserve the normalization of the state it acts on. If U is the matrix representation of a single qubit gate, it is unitary if

$$\mathrm{U}^\dagger \mathrm{U} = \mathrm{U}\mathrm{U}^\dagger = \mathrm{I},$$

where $\mathrm{U}^\dagger$ is the Hermitian adjoint of U, which is obtained by first complex conjugating and then transposing U. Since the unitarity of the gate operator is the only constraint for a single qubit gate, there are essentially infinite number of possible gates. However, any single qubit gate can be decomposed into rotations around two axes of the Bloch sphere [2]. Any single qubit gate can thus be constructed if these rotations can be done somehow.

2.3 Singlet-triplet qubits

The theory of quantum computation is not tied to any particular physical realization of a qubit. In principle, any quantum two-level system, i.e. a system with two available states, is suitable. As a consequence, many possible realizations have emerged, such as ones based on polarization of photons, states of trapped atoms and charge, flux or phase of superconducting circuits [10]. The qubits that are studied in this thesis are based on spin of an electron that is trapped in a quantum dot. The spin of an electron in a quantum dot was proposed as a possible realization of a qubit by Loss and DiVincenzo in 1998 [13]. The singlet-triplet qubits based on two spins in a double quantum dot have been under research recently, because

they offer some benefits compared to the one spin versions, and therefore, they are studied in this thesis.

2.3.1 Decoherence and the optimality of singlet-triplet qubits

One of the greatest challenges in quantum computation is the handling of quantum noise, also known as decoherence, which is caused by the coupling between a quantum computer and its environment [2]. Fortunately, there are quantum error correction algorithms that allow arbitrarily long quantum computations to be carried out if the ratio of the gate operation time and the decoherence time is within some tolerance [14]. Gate operation time is the time needed to perform one quantum gate operation, i.e. a manipulation of one or multiple qubits, and decoherence time is the time over which the phase of a superposition of qubit basis states is well defined [15].

In GaAs based quantum dots, a major cause of decoherence is the hyperfine interaction between the spin of an electron in a dot and the spins of the surrounding Ga and As nuclei [16]. To protect the qubits from the hyperfine interaction, Levy proposed the singlet and triplet states of the spins of two electrons in a double quantum dot as another possible realization of qubit basis states [17]. Singlet-triplet qubits have natural protection for the decoherence due to the hyperfine interaction. Choosing $|S\rangle$ and $|T_0\rangle$, the spin eigenstates of the two-electron system for which $\langle S_z\rangle = 0$, as the logical basis of the qubit makes it immune to uniform fluctuations of the magnetic field [18]. The notation $\langle S_z\rangle$ means the expectation value of the spin operator S_z. In addition, decoherence due to spatial variations in the hyperfine field can be cancelled for the most part by performing a spin echo pulse sequence [11, 19].

2.3.2 Eigenstates of two-spin system

In the so called uncoupled representation, the system of two spins is presented using the eigenstates of the single-electron S_z operators S_{1z} and S_{2z}. The spin states in this representation are thus $|\downarrow\downarrow\rangle$, $|\uparrow\downarrow\rangle$, $|\downarrow\uparrow\rangle$ and $|\uparrow\uparrow\rangle$. The arrows show the alignment of the spins one and two respectively, and thus define the eigenvalues of the S_{1z} and S_{2z} operators. The up spin corresponds to the quantum number $m_S = 1/2$ and the down spin to $m_S = -1/2$. If $|m_S\rangle$ is an eigenstate of S_z with a quantum number m_S, it follows the eigenvalue equation $S_z|m_S\rangle = \hbar m_S|m_S\rangle$. Therefore, the eigenvalue equations for the state $|\downarrow\uparrow\rangle$, for example, are $S_{1z}|\downarrow\uparrow\rangle = -\frac{\hbar}{2}|\downarrow\uparrow\rangle$ and $S_{2z}|\downarrow\uparrow\rangle = \frac{\hbar}{2}|\downarrow\uparrow\rangle$.

If the Hamiltonian of the two-electron system commutes with a chosen set of spin operators, the Hamiltonian and the spin operators have a common set of eigenstates. This also implies that the total wave function of the system is separable to a product of spatial and spin parts and can be written [20]

$$\Psi(\mathbf{r}_1, \mathbf{S}_1; \mathbf{r}_2, \mathbf{S}_2) = \varphi(\mathbf{r}_1, \mathbf{r}_2)\chi(\mathbf{S}_1, \mathbf{S}_2).$$

Since electrons are fermions, the total wave function of two electrons must be antisymmetric with respect to the exchange of the electrons. This property is common to all many-body systems of fermions [21]. The only way to make the total wave function antisymmetric is to have a symmetric spatial wave function and an antisymmetric spin wave function or the other way around.

As the single-electron Hamiltonian (2.5) contains spin operator in the Zeeman term, it does not commute with a general spin operator. However, since the magnetic field is chosen to have only z-component, only the S_z part of the spin operator remains, and the Hamiltonian obviously commutes with the single-electron spin operators S_{1z} and S_{2z}. Although these spin operators commute with the Hamiltonian, the spin states $|\uparrow\downarrow\rangle$ and $|\downarrow\uparrow\rangle$ in the uncoupled representation are not properly symmetrized with respect to exchange of the spins and thus they cannot be proper two-electron states. The solution is to switch to the coupled representation of the spin states.

In the coupled representation, the state of a two-spin system is presented using the eigenstates of the spin operators S^2 and S_z, which are of the form

$$\begin{aligned} S^2 &= (\mathbf{S}_1 + \mathbf{S}_2)^2 = S_1^2 + S_2^2 + 2\mathbf{S}_1 \cdot \mathbf{S}_2 \\ S_z &= S_{1z} + S_{2z}. \end{aligned} \tag{2.10}$$

These spin operators also commute with the Hamiltonian because of the commutator relation $[S_z, S^2] = S_z S^2 - S^2 S_z = 0$, and they have four common eigenstates, a singlet state $|S\rangle$ and three triplet states $|T_-\rangle$, $|T_0\rangle$ and $|T_+\rangle$. These states can be represented as linear combinations of the eigenstates of the single-electron operators S_{1z} and S_{2z}, and they become

$$\begin{aligned} |S\rangle &= \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) \\ |T_-\rangle &= |\downarrow\downarrow\rangle \\ |T_0\rangle &= \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle) \\ |T_+\rangle &= |\uparrow\uparrow\rangle. \end{aligned} \tag{2.11}$$

The spin quantum number of the singlet state is $s = 0$ and that of the triplet states is $s = 1$. Another difference between the singlet state and the triplet states is their exchange symmetry. If P_{12} is a permutation operator that exchanges the particles 1 and 2, $|S\rangle$ is an eigenstate of P_{12} with eigenvalue equation

$$P_{12}|S\rangle = \frac{1}{\sqrt{2}}(|\downarrow\uparrow\rangle - |\uparrow\downarrow\rangle) = -\frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) = -|S\rangle.$$

The singlet state is thus antisymmetric with respect to the exchange of the spins, whereas the triplet states are symmetric. The quantum numbers corresponding to the S_z operator are $m_S = 0$ for $|S\rangle$ and $|T_0\rangle$, $m_S = -1$ for $|T_-\rangle$ and $m_S = 1$ for $|T_+\rangle$.

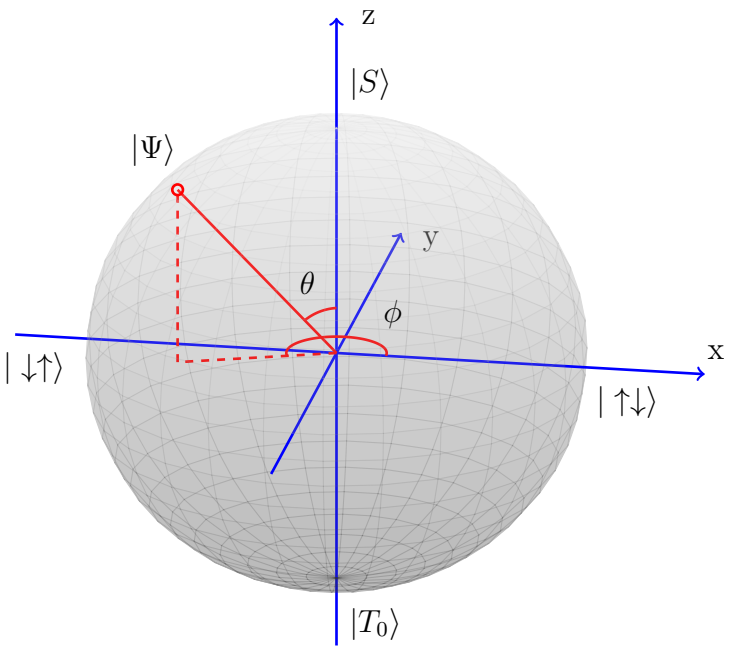

Figure 2.2: Representation of the singlet-triplet qubit state in the Bloch sphere.

2.3.3 Single qubit control

As was noted earlier, the optimal basis states for a singlet-triplet qubit are the singlet state $|S\rangle$, obviously, and the triplet state $|T_0\rangle$. Since $m_S = 0$ for both of these states, the expectation value of the S_z operator is $\langle S_z\rangle = 0$, which makes the $|S\rangle$ and $|T_0\rangle$ states insensitive to uniform fluctuations of the magnetic field. The state of the singlet-triplet qubit can be written in form

$$|\psi\rangle = \cos\frac{\theta}{2}|S\rangle + e^{i\phi}\sin\frac{\theta}{2}|T_0\rangle, \tag{2.12}$$

which is the same as Eq. (2.9) with substitutions $|0\rangle \rightarrow |S\rangle$ and $|1\rangle \rightarrow |T_0\rangle$. The global phase factor has also been omitted because it has no physical meaning. This state can be represented as a point in the Bloch sphere shown in Fig. 2.2. The states at the intersections of the sphere and the x-axis have been calculated by plugging in the corresponding angles, $\phi = 0$, $\theta = \pi/2$ and $\phi = \pi$, $\theta = \pi/2$, to the angle representation of the state and writing $|S\rangle$ and $|T_0\rangle$ as in (2.11).

In the earlier section on quantum computing, it was noted that any single qubit gate can be decomposed into rotations around two axes of the Bloch sphere. Rotations around the z-axis ($|S\rangle - |T_0\rangle$) are driven by the exchange splitting J between the singlet and triplet states. The exchange splitting is the energy difference between the lowest singlet and triplet states,

$$J = E_{T_0} - E_S, \tag{2.13}$$

which is caused by the exchange interaction between the two electrons. The exchange interaction is a purely quantum mechanical effect that affects indistinguishable particles, such as electrons. The two-electron system was shown to have four possible

spin eigenstates, an antisymmetric singlet state and three symmetric triplet states. Since the total wave function of indistinguishable fermions must be antisymmetric, the spatial wave functions corresponding to the singlet state must be symmetric and that of the triplet state must be antisymmetric. The spatial wave function that is antisymmetric with respect to the interchange of the position coordinates of the electrons must be zero if the electrons are at the same position. Therefore, the electrons behave as if they repel each other, and they are statistically further away from each other. For the symmetric wave function, the effect is opposite. As a consequence, the singlet state has a higher probability density between the two quantum dots than the triplet states. This, in turn, has two opposite effects on the energies of the states. On one hand, Coulomb repulsion increases and raises the total energy, and on the other hand, the electron density is more centred to the area with low external potential, which lowers the total energy. In the GaAs double quantum dot system that is studied in this thesis, the second effect overcomes the first one, and thus the singlet energy is always lower than the triplet energy (assuming that the global magnetic field is zero).

The exchange splitting is sensitive to the distance and barrier height between the two parabolic quantum dots as well as to the detuning of the dots, i.e. to the energy difference ε between the minima of the two dots. Since detuning is the easiest to control, it is used to turn on the exchange splitting and, in consequence, the rotation around the z-axis of the Bloch sphere [11]. The effect of detuning on the exchange splitting is based on the mixing of the two-electron states $|(1,1)S\rangle$ and $|(2,0)S\rangle$, where $(1,1)$ denotes the spatial part of the wave function where both dots are occupied by one electron and $(2,0)$ denotes the wave function where both electrons are in the same dot (the left one in this case). When the potential minimum of one dot is lowered with respect to the other, the singlet state allows the mixing of charge states $(1,1)$ and $(2,0)$, which offers lower total energy than a pure $(1,1)$ state. The triplet state needs a substantially higher detuning than the singlet state before the charge state $(2,0)$ becomes available because the antisymmetric spatial wave function prevents the electrons from being in a same position, which results in an effective mutual repulsion of the electrons.

Rotation of the qubit state around the x-axis of the Bloch sphere ($|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle$) can be driven by the z component of a magnetic field gradient between the two quantum dots. The magnetic field can be generated using dynamic nuclear polarization, i.e. by transferring spin from the electronic system to the underlying Ga and As nuclei [22]. The combined axis of rotation can be written

$$\hat{\mathbf{n}} = J\hat{\mathbf{e}}_{\mathbf{z}} + g^*\mu_B\Delta B_z\hat{\mathbf{e}}_{\mathbf{x}} \tag{2.14}$$

and the rotation frequency is given by

$$f = \sqrt{J^2 + (g^*\mu_B\Delta B_z)^2}/h, \tag{2.15}$$

where ΔB_z is the z component of the magnetic field difference between the dots [22].

Chapter 3

Computational methods

3.1 Lagrange mesh method

Lagrange mesh method is used in this work to solve the Schrödinger equation with the Hamiltonian (2.5). The Lagrange mesh method is an approximately variational calculation method that resembles a mesh calculation. This means that, in contrast to traditional variational method, no integrals need to be calculated. The integrals are approximated by a Gauss quadrature instead, which makes the Lagrange mesh method extremely versatile, as almost arbitrary potentials can be handled. The mesh calculation properties are achieved by using a basis of Lagrange functions [23].

3.1.1 Lagrange basis

A set of N Lagrange functions L_k defined over an interval (a, b) is associated with N mesh points $x_k \in (a, b)$. A Gauss quadrature

$$\int_a^b dx L_i(x) \approx \sum_{k=1}^{N} \lambda_k L_i(x_k) \tag{3.1}$$

is associated to the mesh. Lagrange functions are infinitely differentiable functions which satisfy Lagrange conditions

$$L_i(x_j) = \lambda_i^{-1/2} \delta_{ij}. \tag{3.2}$$

Additionally, the Gauss quadrature for a product of Lagrange functions $L_i(x)L_j(s)$ is exact. From the exactness of the Gauss quadrature and the conditions (3.2), it follows that the Lagrange functions are orthonormal

$$\int_a^b dx L_i(x) L_j(x) = \sum_{k=1}^{N} \lambda_k L_i(x_k) L_j(x_k) = \delta_{ij}. \tag{3.3}$$

There are many possible functions that can be used in Lagrange mesh calculations. In this work, one of the most simple ones, a sinc function, is used. The sinc mesh is defined over the interval $(-\infty, \infty)$, but it treats only fairly well localized

wave functions. Sinc functions are used as the Lagrange functions in this thesis because they have the exponentially decaying asymptotic behaviour that the electron states are expected to have [24]. The mesh points are distributed uniformly around the origin

$$x_a = a, \quad a \in \left\{ -\frac{N-1}{2}, -\frac{N-1}{2}+1, \ \ldots \ , \frac{N-1}{2} \right\}, \tag{3.4}$$

and the weights in the Gauss quadrature are $\lambda_a = 1$. The Lagrange-sinc functions are defined as

$$L_a(x) = \mathrm{sinc}(x-a) = \frac{\sin[\pi(x-a)]}{\pi(x-a)}. \tag{3.5}$$

For the quantum dot calculations, a two-dimensional Lagrange grid is needed. The one-dimensional grid is easily generalized to two dimensions as

$$L_{\mathbf{a}}(x,y) = L_{a_x}(x)L_{a_y}(y) = \mathrm{sinc}(x-a_x)\,\mathrm{sinc}(y-a_y). \tag{3.6}$$

If the minimum of the potential is inside the area $(-\frac{L}{2}, \frac{L}{2}) \times (-\frac{L}{2}, \frac{L}{2})$, the wave function of the electron tends to zero outside it. When the area is chosen to be large enough, the wave function outside the square nearly vanishes. Now the grid points, where the Lagrange functions are centered at, should be scaled so that they are located inside the chosen area. The scaled mesh points are

$$\mathbf{r_a} = \frac{L}{N}\mathbf{a}, \quad a_x, a_y \in \left\{ -\frac{N-1}{2}, -\frac{N-1}{2}+1, \ \ldots \ , \frac{N-1}{2} \right\}, \tag{3.7}$$

where N is the number of mesh points in one dimension. The scaled two-dimensional Lagrange functions are

$$L_{\mathbf{a}}(\mathbf{r}) = \frac{N}{L}\,\mathrm{sinc}\left[\frac{N}{L}(x-x_{a_x})\right]\mathrm{sinc}\left[\frac{N}{L}(y-y_{a_y})\right]. \tag{3.8}$$

These scaled functions satisfy the Lagrange conditions, if the weights are chosen to be $\lambda_{\mathbf{a}} = h^2$, where $h = L/N$ is the grid spacing.

3.1.2 Hamiltonian matrix elements

Using the derived Hamiltonian (2.5), the corresponding time-independent Schrödinger equation can be written as

$$\left(-\frac{1}{2}\nabla^2 + V(\mathbf{r}) - i\frac{\omega_c}{2}(y\partial_x - x\partial_y)\right)\psi(\mathbf{r}) = E\psi(\mathbf{r}), \tag{3.9}$$

where the external potential and the parabolic term from the magnetic field have been combined to $V(\mathbf{r})$. A variational trial wave function is expanded in the orthonormal basis of Lagrange functions

$$\psi(\mathbf{r}) = \sum_{\mathbf{a}} c_{\mathbf{a}} L_{\mathbf{a}}(\mathbf{r}). \tag{3.10}$$

Substituting the trial wave function to the Schrödinger equation, multiplying by $L_{\mathbf{a}'}(\mathbf{r})$ and integrating over the whole domain leads to a set of variational equations

$$\sum_{\mathbf{a}}(T_{\mathbf{a}'\mathbf{a}} + V_{\mathbf{a}'\mathbf{a}} + B_{\mathbf{a}'\mathbf{a}})c_{\mathbf{a}} = Ec_{\mathbf{a}'}. \tag{3.11}$$

The potential energy matrix elements are approximated using the Gauss quadrature (3.1)

$$V_{\mathbf{a}'\mathbf{a}} \approx V(\mathbf{r}_{\mathbf{a}})\delta_{a'_x a_x}\delta_{a'_y a_y}. \tag{3.12}$$

The kinetic energy matrix elements are

$$T_{\mathbf{a}'\mathbf{a}} = -\frac{1}{2}\int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy\; L_{\mathbf{a}'}(\mathbf{r})\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) L_{\mathbf{a}}(\mathbf{r}) \tag{3.13}$$

and the magnetic field induced matrix elements $B_{\mathbf{a}'\mathbf{a}}$ are

$$B_{\mathbf{a}'\mathbf{a}} = -i\frac{\omega_c}{2}\int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy\; L_{\mathbf{a}'}(\mathbf{r})(x\partial_y - y\partial_x)L_{\mathbf{a}}(\mathbf{r})\,. \tag{3.14}$$

The matrix elements $T_{\mathbf{a}'\mathbf{a}}$ and $B_{\mathbf{a}'\mathbf{a}}$ can be both calculated analytically using contour integration techniques in the complex plane, but those calculations are not shown here. After calculating the kinetic energy matrix elements $T_{\mathbf{a}'\mathbf{a}}$ and the magnetic field induced elements $B_{\mathbf{a}'\mathbf{a}}$, the complete Hamiltonian matrix elements become [24, 25]

$$H_{a'_x a'_y a_x a_y} = \begin{cases} \frac{\pi^2}{3h^2} + V(\mathbf{r}_{\mathbf{a}}) & ,\; a'_x = a_x, a'_y = a_y \\ (-1)^{a'_y - a_y}\left[\frac{1}{h^2(a'_y - a_y)^2} + \frac{i\omega_c}{2}\frac{a_x}{a'_y - a_y}\right] & ,\; a'_x = a_x, a'_y \neq a_y \\ (-1)^{a'_x - a_x}\left[\frac{1}{h^2(a'_x - a_x)^2} - \frac{i\omega_c}{2}\frac{a_y}{a'_x - a_x}\right] & ,\; a'_x \neq a_x, a'_y = a_y \\ 0 & ,\; a'_x \neq a_x, a'_y \neq a_y. \end{cases}$$

Diagonalizing this Hamiltonian matrix gives the single-particle states and their energies. The global magnetic field is not used in any part of this thesis, and therefore, $\omega_c = 0$ and the imaginary part of the Hamiltonian disappears.

3.2 Exact diagonalization

The Lagrange mesh method provides the eigenstates of a single electron in a quantum dot system, and these single-electron states are used as a basis for constructing the many-electron states. The quantum many-body problem is computationally demanding because the size of the many-body basis grows extremely fast as a function of the number of single-electron states and the number of particles. If one takes N_s single-electron states and builds a basis for N electrons, the size of this basis is $\binom{N_s}{N}$. When building the many-body basis, the proper symmetrization of the many-body states must be taken into account.

If the number of electrons is sufficiently small, it is possible to build the many-body basis and solve the ground state of the system without making any simplifying approximations. This approach is called exact diagonalization method. The exact diagonalization technique that is used in this thesis relies on the second quantization formalism of quantum mechanics, where the many-body basis states are represented as occupations of the single-particle states, and all operators are written using creation and annihilation operators that create or destroy a particle in a certain single-particle state. The second quantization Hamiltonian matrix is simple to create on a computer, and the result is a huge, but sparse, matrix. The lowest eigenvalue and the corresponding eigenvector of the Hamiltonian matrix, i.e. the ground state, can be solved efficiently using the Lanczos diagonalization algorithm.

3.2.1 Second quantization

Since we are dealing with electrons, which are fermions, the many-body eigenstates must be antisymmetric. In the case of no interactions, the correct antisymmetrization can be achieved by writing the N-body states using Slater determinants

$$\Psi(\xi_1, \xi_2, ..., \xi_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_{s_1}(\xi_1) & \psi_{s_2}(\xi_1) & \cdots & \psi_{s_N}(\xi_1) \\ \psi_{s_1}(\xi_2) & \psi_{s_2}(\xi_2) & \cdots & \psi_{s_N}(\xi_2) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_{s_1}(\xi_N) & \psi_{s_2}(\xi_N) & \cdots & \psi_{s_N}(\xi_N) \end{vmatrix}. \tag{3.15}$$

Here $\xi_i = (\mathbf{r}_i, \sigma_i)$ hold both the space coordinates and the spin coordinate of a particle, and $S = \{s_1, s_2, \ldots, s_N\} \subseteq \{1, 2, \ldots, N_s\}$ is a subset of single-particle state indices. The Slater determinant gives all $N!$ possible permutations of particles and the correct sign for each term. To get all possible many-body states, one must consider all N-combinations of the set of N_s single-particle states. The Slater determinant is not a correct form of a wave function when interaction between particles exists, but the set of determinants with all possible combinations of single-electron states can be used as a many-particle basis in a variational calculation. Slater determinant formalism is notationally and computationally heavy though, and therefore, a different approach is taken.

A convenient notation in the presence of interactions is the so called second quantization. In the second quantization formalism, many-body states are represented by occupation numbers of each single-particle state, and each many-body state is characterized by a sequence of creation operators, which create the particles to the single-particle states. Thus, we can write a general N-particle state as

$$|n_1 n_2 \ldots\rangle = (a_1^\dagger)^{n_1} (a_2^\dagger)^{n_2} \ldots |0\rangle, \tag{3.16}$$

where $a_i^\dagger$ is the creation operator of the i'th single-particle state, and $|0\rangle = |00\ldots\rangle$ is the vacuum state with no particles. The bra-states are created similarly but using the annihilation operators,

$$\langle \ldots n_2 n_1| = \langle 0| \ldots (a_2)^{n_2} (a_1)^{n_1}. \tag{3.17}$$

The annihilation operators have of course the property $a_i|0\rangle = 0$. The possible occupation numbers for fermions are either 0 or 1, because of the Pauli exclusion principle. The occupation number states are orthogonal and normalized, and thus fulfill

$$\langle \dots m_2 m_1 | n_1 n_2 \dots \rangle = \delta_{m_1 n_1} \delta_{m_2 n_2} \dots . \tag{3.18}$$

The Fermionic annihilation and creation operators satisfy the anticommutation relations

$$[a_j^\dagger, a_k^\dagger]_+ = [a_j, a_k]_+ = 0, \qquad [a_j^\dagger, a_k]_+ = \delta_{jk}, \tag{3.19}$$

where $[a, b]_+ = ab + ba$. The anticommutation relations guarantee the correct antisymmetrization of the states. The Hamiltonian (2.7) is expressed completely using the creation and annihilation operators, which results in [21]

$$H = \sum_{j,k} h_{jk} a_j^\dagger a_k + \frac{1}{2} \sum_{i,j,k,l} U_{ijkl} a_i^\dagger a_j^\dagger a_k a_l. \tag{3.20}$$

Here

$$h_{jk} = \langle \psi_j | H_0 | \psi_k \rangle \tag{3.21}$$

are the matrix elements of the single-particle Hamiltonian and

$$U_{ijkl} = \langle \psi_i \psi_j | U | \psi_k \psi_l \rangle \tag{3.22}$$

are the matrix elements of the interaction term in the single-particle basis.

3.2.2 Interaction matrix elements

The calculation of the interaction matrix elements U_{ijkl} of Eq. (3.22) is quite tricky, because of the $1/r$ form of the interaction. The following analytical manipulations are needed to calculate those elements [25].

Let ψ_i be the single-particle eigenfunctions expanded in the sinc basis of Eq. (3.8),

$$\psi_i(\mathbf{r}) = \sum_{\mathbf{a}} \alpha_{\mathbf{a}}^i L_{\mathbf{a}}(\mathbf{r}) \ .$$

The interaction matrix elements are then

$$\begin{aligned} V_{ijkl} &= \int_{\mathbb{R}^2} \mathrm{d}\,\mathbf{r}_1 \int_{\mathbb{R}^2} \mathrm{d}\,\mathbf{r}_2 \Psi_i^*(\mathbf{r}_1) \Psi_j^*(\mathbf{r}_2) \frac{1}{r_{12}} \Psi_k(\mathbf{r}_1) \Psi_l(\mathbf{r}_2) \\ &= \sum_{\mathbf{a},\mathbf{b},\mathbf{c},\mathbf{d}} \alpha_{\mathbf{a}}^{i*} \alpha_{\mathbf{b}}^{j*} \alpha_{\mathbf{c}}^k \alpha_{\mathbf{d}}^l \int_{\mathbb{R}^2} \mathrm{d}\,\mathbf{r}_1 \int_{\mathbb{R}^2} \mathrm{d}\,\mathbf{r}_2 \\ &\quad \times L_{\mathbf{a}}(\mathbf{r}_1) L_{\mathbf{b}}(\mathbf{r}_2) \frac{1}{r_{12}} L_{\mathbf{c}}(\mathbf{r}_1) L_{\mathbf{d}}(\mathbf{r}_2) \\ &= \sum_{\mathbf{a},\mathbf{b},\mathbf{c},\mathbf{d}} \alpha_{\mathbf{a}}^{i*} \alpha_{\mathbf{b}}^{j*} \alpha_{\mathbf{c}}^k \alpha_{\mathbf{d}}^l v_{\mathbf{abcd}} \ , \end{aligned} \tag{3.23}$$

where the expansion coefficients α multiply the interaction matrix elements $v_{\mathbf{abcd}}$ between the sinc basis functions. To calculate these, we start with the two-dimensional Fourier transform of $1/r_{12}$, namely

$$\begin{aligned}\mathcal{F}\left[\frac{1}{r_{12}}\right](\mathbf{k}) &= \int_{\mathbb{R}^2} \mathrm{d}\,\mathbf{r}_{12} \frac{e^{-i\mathbf{k}\cdot\mathbf{r}_{12}}}{r_{12}} \\ &= \int_0^\infty \mathrm{d}\,r_{12} \int_0^{2\pi} \mathrm{d}\,\theta e^{ikr_{12}\cos(\theta+\pi)}\,,\end{aligned} \tag{3.24}$$

where θ is the angle between $\mathbf{r}_{12}$ and $\mathbf{k}$. Using the Jacobi-Anger identity of Bessel functions,

$$e^{iz\cos\phi} = \sum_{n=-\infty}^{\infty} i^n J_n(z) e^{in\phi}\,,$$

leads to

$$\begin{aligned}\mathcal{F}\left[\frac{1}{r_{12}}\right](\mathbf{k}) &= \int_0^\infty \mathrm{d}\,r_{12} \int_0^{2\pi} \mathrm{d}\,\theta \sum_{n=-\infty}^{\infty} (-i)^n J_n(kr_{12}) e^{in\theta} \\ &= 2\pi \int_0^\infty \mathrm{d}\,r_{12} J_0(kr_{12}) = \frac{2\pi}{k}\,.\end{aligned}$$

The potential $1/r_{12}$ can now be written as the inverse Fourier transform of $\mathcal{F}\left[\frac{1}{r_{12}}\right]$ as:

$$\begin{aligned}\frac{1}{r_{12}} &= \left(\mathcal{F}^{-1}\circ\mathcal{F}\right)\left[\frac{1}{r_{12}}\right] = \frac{1}{(2\pi)^2}\int_{\mathbb{R}^2} \mathrm{d}^2\mathbf{k}\frac{2\pi}{k} e^{i\mathbf{k}\cdot\mathbf{r}_{12}} \\ &= \frac{1}{2\pi}\int_{\mathbb{R}^2} \mathrm{d}\,k_x\,\mathrm{d}\,k_y\,\frac{1}{k}\,e^{ik_x(x_2-x_1)}e^{ik_y(y_2-y_1)}\,.\end{aligned} \tag{3.25}$$

With the identity of Eq. (3.25), the integrations over different coordinates factorize in the interaction matrix element:

$$\begin{aligned}v_{\mathbf{abcd}} = \frac{1}{2\pi h}\int_{\mathbb{R}^2} \mathrm{d}\,\mathbf{k}\,\frac{1}{k} \\ \times \int_{-\infty}^{\infty} \mathrm{d}\,x_1\,\mathrm{sinc}(x_1-a_x)\,\mathrm{sinc}(x_1-c_x)e^{ik_x x_1} \\ \times \int_{-\infty}^{\infty} \mathrm{d}\,y_1\,\mathrm{sinc}(y_1-a_y)\,\mathrm{sinc}(y_1-c_y)e^{ik_y y_1} \\ \times \int_{-\infty}^{\infty} \mathrm{d}\,x_2\,\mathrm{sinc}(x_2-b_x)\,\mathrm{sinc}(x_2-d_x)e^{-ik_x x_2} \\ \times \int_{-\infty}^{\infty} \mathrm{d}\,y_2\,\mathrm{sinc}(y_2-b_y)\,\mathrm{sinc}(y_2-d_y)e^{-ik_y y_2}.\end{aligned} \tag{3.26}$$

The sinc functions can be replaced by their integral representation

$$\mathrm{sinc}(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{d}\,t\; e^{ixt}\,,$$

and the integrals over x and y coordinates are of the form

$$I_{ab}(k) = \int_{-\infty}^{\infty} \mathrm{d}\, x \,\mathrm{sinc}(x-a)\,\mathrm{sinc}(x-b)e^{ikx}$$
$$= \begin{cases} \frac{i\,\mathrm{sign}(k)}{2\pi(a-b)}(-1)^{a-b}\left(e^{ika}-e^{ikb}\right) & , \ |k| \leq 2\pi, \ a \neq b \\ \frac{1}{2\pi}e^{ika}(2\pi - |k|) & , \ |k| \leq 2\pi, \ a = b \\ 0 & , \ |k| > 2\pi\,. \end{cases}$$

By substituting this result into Eq. (3.26), the original four-dimensional integral over two planes reduces into a two-dimensional integral over a finite square in k-space,

$$\begin{aligned} v_{\mathbf{abcd}} &= \frac{1}{2\pi h}\int_{-2\pi}^{2\pi}\mathrm{d}\,k_x \int_{-2\pi}^{2\pi}\mathrm{d}\,k_y \frac{1}{k} \\ &\quad \times I_{a_x c_x}(k_x) I_{a_y c_y}(k_y) I_{b_x d_x}(k_x) I_{b_y d_y}(k_y) \\ &= \frac{1}{2\pi h}\int_0^{2\pi}\mathrm{d}\,\theta \int_0^{K(\theta)}\mathrm{d}\,k \\ &\quad \times I_{a_x c_x}(-k\cos(\theta)) I_{a_y c_y}(-k\sin(\theta)) \\ &\quad \times I_{b_x d_x}(k\cos(\theta)) I_{b_y d_y}(k\sin(\theta))\ , \end{aligned} \tag{3.27}$$

where $K(\theta) = 2\pi/\max(|\cos(\theta)|, |\sin(\theta)|)$ is the radial integration limit corresponding to the square. The last form can be used in numerical calculations.

One can see that in Eq. (3.27), one obtains five different integrals depending on how many of the four functions I_{ab} have the same indices. In addition, the case with two equal index pairs is naturally split into two cases, depending on whether the equal indices belong to the same Cartesian component of k. In most cases, some further analytic work can be done to handle the angular integral. For instance, in the case when all the index pairs differ, such that $a_x \neq c_x$, $b_x \neq d_x$, $a_y \neq c_y$ and $b_y \neq d_y$, the integrand can be written as a sum of terms of the form $\cos\{k[m\cos(\theta) + n\sin(\theta)]\}$, and the angular part can be integrated analytically, and we are left with a one-dimensional numerical integral. In this way, we are able to calculate the interaction matrix elements between the sinc basis functions, and then for any external confinement potential, Eq. (3.23) can be used to construct V_{ijkl}.

3.2.3 Many-particle basis

The basis for many-particle states is formed using the calculated single-particle states. The basis states are written in the occupation number formalism as explained in the second quantization section. Since the basis in computations must be finite, a set of N_s lowest energy single-particle states is used. Therefore, the basis states for N electrons are

$$|\mathbf{n}\rangle = |n_1 n_2 \ldots n_{N_s}\rangle, \qquad \sum_{k=1}^{N_s} n_k = N \tag{3.28}$$

and the size of the basis is

$$N_{mb} = \binom{N_s}{N} = \frac{N_s!}{N!(N_s - N)!}.$$

A trial wave function in this basis can be written as

$$|\Psi\rangle = \sum_{\mathbf{n}} \alpha_{\mathbf{n}} |\mathbf{n}\rangle. \tag{3.29}$$

When the occupation number states are written in the basis of the eigenstates of the single-particle Hamiltonian, the single-particle part of the second quantization Hamiltonian is diagonal

$$H = \sum_j E_j a_j^\dagger a_j + \frac{1}{2} \sum_{i,j,k,l} U_{ijkl} a_i^\dagger a_j^\dagger a_k a_l, \tag{3.30}$$

and E_j are the single-particle energies. The time independent Schrödinger equation is written using the trial wave function

$$H \sum_{\mathbf{n}} \alpha_{\mathbf{n}} |\mathbf{n}\rangle = E \sum_{\mathbf{n}} \alpha_{\mathbf{n}} |\mathbf{n}\rangle. \tag{3.31}$$

Operating on this equation by $\langle \mathbf{m}|$ leads to a set of equations

$$\sum_{\mathbf{n}} H_{\mathbf{mn}} \alpha_{\mathbf{n}} = E \alpha_{\mathbf{m}}, \tag{3.32}$$

where

$$H_{\mathbf{mn}} = \langle \mathbf{m}|H|\mathbf{n}\rangle.$$

The matrix form of these equations is

$$\mathrm{H}\,\boldsymbol{\alpha} = E\,\boldsymbol{\alpha}\,, \tag{3.33}$$

which is the eigenvalue equation of the many-body Hamiltonian matrix H. The coefficients $\alpha_{\mathbf{n}}$, which define the many-particle states, can thus be found by diagonalizing H. However, the size of the basis can be enormous, and thus the diagonalization of the matrix by standard means may be impractical or even impossible. Since the many-body Hamiltonian matrix is sparse, i.e. most of the elements are zero, iterative Lanczos diagonalization algorithm offers effective means to calculate the ground state of the many-particle system.

Different spins of particles must be taken into account in the exact diagonalization calculation, because particles with the same spin cannot occupy the same single-particle state, but particles with different spins can. Thus, there can be two particles in the same single-particle state if the spins of the particles are opposite. This can be handled by treating particles with different spins in separate components, i.e. by using two separate sets of single-particle states.

3.2.4 Dynamics

The exact diagonalization technique gives the ground state of a many-body system, but the interesting part is often the dynamics of the system, i.e. the time evolution of the ground state. The time evolution of a quantum mechanical state is given by the time dependent Schrödinger equation

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle = H(t)|\psi(t)\rangle. \tag{3.34}$$

Assuming that the Hamiltonian of the system is time-independent, the solution of this differential equation is simply

$$|\psi(t)\rangle = e^{-\frac{i}{\hbar}Ht}|\psi(0)\rangle. \tag{3.35}$$

If the Hamiltonian is time-dependent, a good approximation for the time evolution is obtained by discretizing the time to steps of Δt. Assuming that the Hamiltonian is approximately constant during time Δt, at time $t + \Delta t$ the state is given by

$$|\psi(t + \Delta t)\rangle \approx e^{-\frac{i}{\hbar}H(t)\Delta t}|\psi(t)\rangle. \tag{3.36}$$

In practise, the initial many-body state $|\psi(0)\rangle$ is represented by the eigenvector $\boldsymbol{\alpha}_0$ that corresponds to the lowest eigenvalue of the many-body Hamiltonian matrix in Eq. (3.33). More generally, the vector representation of the state at time t is

$$|\psi(t)\rangle \to \boldsymbol{\alpha}_t = \begin{bmatrix} \alpha_{t,1} \\ \alpha_{t,2} \\ \vdots \\ \alpha_{t,N_{mb}} \end{bmatrix}, \tag{3.37}$$

where $\alpha_{t,i}$ are the coefficients of the many-body state expansion in Eq. (3.29). In this matrix representation, the time evolution of the state given by

$$\boldsymbol{\alpha}_{t+\Delta t} = e^{-i\Delta t\,\mathrm{H}_t}\boldsymbol{\alpha}_t\,, \tag{3.38}$$

where H_t is the Hamiltonian matrix at time t and $\hbar$ has been removed due to the use of effective atomic units. Calculation of the matrix exponential $e^{-i\Delta t\,\mathrm{H}_t}$ poses a serious issue because H_t is typically huge. Since full diagonalization of H_t is impractically slow, the standard way of calculating the matrix exponential through the diagonal form of the matrix is impossible. An effective algorithm for the matrix exponential problem is presented in section 3.3.2.

3.3 Algorithms for large, sparse matrices

3.3.1 Lanczos diagonalization

The idea of the Lanczos algorithm [26] is to map the original enormous eigenvalue problem to a much smaller tridiagonal problem. This is done using Krylov subspaces

as a basis. The original matrix must be Hermitian, which should of course be true for a Hamiltonian matrix. A Krylov subspace of dimension m is defined by

$$\mathcal{K}^m(\mathbf{b}, \mathrm{H}) = \mathrm{span}\{\mathbf{b}, \mathrm{H}\mathbf{b}, \mathrm{H}^2\mathbf{b}, \ldots, \mathrm{H}^{m-1}\mathbf{b}\} \tag{3.39}$$

assuming that $\mathbf{b}$ is not orthogonal to any of the eigenvectors of H. The assumption is realistic if $\mathbf{b}$ is chosen randomly. The basis vectors of the Krylov subspace are orthonormalized using the Gram-Schmidt orthonormalization process to form the Lanczos basis $\{\mathbf{b}_0, ..., \mathbf{b}_{m-1}\}$.

The Gram-Schmidt orthonormalization process of the basis vectors starts by choosing the random vector $\mathbf{b}$ and normalizing it, i.e. $\mathbf{b}_0 = \frac{\mathbf{b}}{|\mathbf{b}|}$. The next basis vector can be solved from the equation

$$\mathrm{H}\mathbf{b}_0 = \alpha_0\mathbf{b}_0 + \beta_1\mathbf{b}_1,$$

where $\alpha_0 = (\mathrm{H}\mathbf{b}_0, \mathbf{b}_0)$ and $\beta_1 = (\mathrm{H}\mathbf{b}_0, \mathbf{b}_1)$. The equation states that the next Krylov subspace basis vector $\mathrm{H}\mathbf{b}_0$ is a linear combination of the previous vector $\mathbf{b}_0$ and a new vector $\mathbf{b}_1$, which is orthogonal to the previous one. The coefficient β_1 is chosen so that $\mathbf{b}_1$ is normalized. The next vector is given by the equation

$$\mathrm{H}\mathbf{b}_1 = \beta_1\mathbf{b}_0 + \alpha_1\mathbf{b}_1 + \beta_2\mathbf{b}_2,$$

where the assumption that the matrix is Hermitian has been used as $\beta_1 = (\mathrm{H}\mathbf{b}_0, \mathbf{b}_1) = (\mathrm{H}\mathbf{b}_1, \mathbf{b}_0)$. After the calculation of j basis vectors, the complete process at step $j+1$ can be written in matrix form

$$\mathrm{HB}_j = \mathrm{B}_j\mathrm{T}_j + \beta_{j+1}\mathbf{b}_{j+1}\mathbf{e}_j^{\mathrm{T}}\,, \tag{3.40}$$

where B_j has the j orthonormal basis vectors as its columns, $\mathbf{e}_j^{\mathrm{T}} = [0\,0\,\cdots\,1]$ and T_j is a tridiagonal matrix,

$$\mathrm{T}_j = \begin{bmatrix} \alpha_0 & \beta_1 & \cdots & 0 \\ \beta_1 & \alpha_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \beta_j \\ 0 & \cdots & \beta_j & \alpha_j \end{bmatrix}. \tag{3.41}$$

If $\mathrm{H} \in \mathbb{C}^{nxn}$, and all n orthonormal basis vectors could be calculated, the B matrix is unitary and T is a square matrix, and therefore, the Hamiltonian matrix could be written in form

$$\mathrm{H} = \mathrm{BTB}^*. \tag{3.42}$$

In this case, the eigenvalues of the tridiagonal matrix T would be exactly the eigenvalues of H.

When n is very large, it becomes impossible to calculate n basis vectors that are all orthogonal to each other because of the numerical errors, and orthogonalizing a large number of vectors would be too time consuming anyway. Fortunately, the eigenvalues of $\mathrm{T_j}$ are good estimates of the eigenvalues of H already when $j \ll n$. This is due to the optimality of the Krylov subspace as the basis [27]. Especially the lowest and highest eigenvalues converge fast, and thus the energy of the ground state can be calculated accurately with a reasonable computational effort.

3.3.2 Matrix exponentiation

A great deal of algorithms have been developed for the matrix exponentiation problem, but there is no universal solution for it [28]. The optimal algorithm depends on various properties of the matrix, such as the size and density, and whether the matrix is Hermitian (or symmetric) or not. The standard way to calculate the matrix exponential $e^{c\,\mathrm{H}}$ of a $n \times n$ Hermitian matrix H is to diagonalize the matrix to get the decomposition

$$\mathrm{H} = \mathrm{U}\Lambda\mathrm{U}^*,$$

where U is a unitary matrix with the orthonormal eigenvectors of H as its columns and the diagonal matrix Λ has the eigenvalues of H as its elements. Since c is just some complex number, the exponential simplifies to

$$e^{c\,\mathrm{H}} = \mathrm{U}e^{c\,\Lambda}\mathrm{U}^*\,, \tag{3.43}$$

where the exponential of the diagonal matrix Λ is obtained simply by exponentiating its diagonal entries, i.e.

$$e^{c\,\Lambda} = \mathrm{diag}(e^{c\lambda_1}, e^{c\lambda_2}, \ldots, e^{c\lambda_n}).$$

However, if the size of the matrix H is so large that the complete diagonalization cannot be made, which is the case for the many-body Hamiltonian matrix in Eq. (3.38), some approximative algorithm is needed.

To calculate the time evolution of a state vector, one needs to calculate the product $e^{-i\Delta t\,\mathrm{H}_t}\boldsymbol{\alpha}_t$, or more generally $e^{c\,\mathrm{H}}\mathbf{b}$. Since H_t is a large but sparse matrix and only the product is wanted, not the matrix exponential itself, Krylov subspace methods are extremely effective [28].

As H is Hermitian, the same Lanczos tridiagonalization algorithm as in the previous section is used to obtain Eq. (3.40) at step $j+1$. It turns out that a good approximation for $e^{c\,\mathrm{H}}\mathbf{b}$ is given by

$$e^{c\,\mathrm{H}}\mathbf{b} \approx |\mathbf{b}|\mathrm{B}_j\, e^{c\,\mathrm{T}_j}\mathbf{e}_1\,, \tag{3.44}$$

where B_j and T_j are as in Eq. (3.40), and $\mathbf{e}_1 = [1\,0\,\cdots\,0]^{\mathrm{T}}$. A derivation and error analysis of the Lanczos exponentiation method are presented by, for example, Saad [29] and Hochbruck and Lubich [30]. Since the approximation (3.44) is typically good already at step $j \ll n$, the $j \times j$ tridiagonal matrix T_j can be readily diagonalized to form $\mathrm{T}_j = \mathrm{U}_j\Lambda_j\mathrm{U}_j^*$. The matrix exponential $e^{c\,\mathrm{T}_j}$ can then be calculated in the same way as Eq. (3.43), and the complete formula for $e^{c\,\mathrm{H}}\mathbf{b}$ becomes

$$e^{c\,\mathrm{H}}\mathbf{b} \approx |\mathbf{b}|\mathrm{B}_j\mathrm{U}_j\, e^{c\,\Lambda_j}\mathrm{U}_j^*\mathbf{e}_1\,. \tag{3.45}$$

Chapter 4

Gauss quadrature approximation

As the Hamiltonian matrix elements due to kinetic energy and magnetic field can be calculated analytically, only the potential matrix elements (3.12) are approximated by the Gauss quadrature (3.1). This approximation is crucial since it allows the potential elements to be calculated simply by evaluating the potential at the mesh points. It is important that the Gauss quadrature approximation is reasonable for any chosen potential. Otherwise, the most essential advantage of the Lagrange mesh method, handling of an arbitrary potential, is lost. Curiously enough, the parabolic potential, which is used to model the quantum dots in this thesis, seems to be problematic to handle for the chosen sinc-mesh. If the potential matrix elements for a parabolic potential are calculated analytically, instead of using the Gauss quadrature, one of the integrals diverges, which renders the elements useless. The matrix elements that one gets using the Gauss quadrature, however, produce highly accurate results.

The effect of the Gauss quadrature approximation to the eigenstates of a Hamiltonian with a parabolic potential is studied in this chapter. A perfectly reasonable cause for the high accuracy of the Gauss quadrature approximation is found, and the analysis also leads to a method for completely eliminating the effect of the approximation. It seems that these results can be easily generalized to higher degree polynomials, but a rigorous proof for that is beyond the scope of this thesis. If one assumes that the Gauss quadrature approximation works well for a polynomial potential of an arbitrary degree, it is easy to argue that the Lagrange mesh method works for any potential, and the convergence rate depends on how well the potential can be approximated by a polynomial.

4.1 Analytical matrix elements of parabolic potential

First, the potential matrix elements are calculated analytically ignoring the Gauss quadrature. The potential energy matrix elements are defined as

$$V_{\mathbf{a}'\mathbf{a}} = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \; L_{\mathbf{a}'}(\mathbf{r}) V(\mathbf{r}) L_{\mathbf{a}}(\mathbf{r}). \tag{4.1}$$

The parabolic potential in effective atomic units is given by $V(\mathbf{r}) = \frac{1}{2}\omega^2 r^2$. When this is plugged in Eq. (4.1) and the definition of the scaled two-dimensional Lagrange functions (3.8) is used, the matrix elements become

$$\begin{aligned} V_{\mathbf{a'a}} &= \frac{1}{2}\omega^2 \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy\; L_{\mathbf{a'}}(\mathbf{r})(x^2+y^2)L_{\mathbf{a}}(\mathbf{r}) \\ &= \frac{1}{2}\omega^2 \left(\frac{1}{h}\int_{-\infty}^{\infty} dx\; x^2\; \mathrm{sinc}[\frac{1}{h}(x-x_{a'_x})]\; \mathrm{sinc}[\frac{1}{h}(x-x_{a_x})]\, \delta_{a'_y a_y} \right. \\ &\quad \left. + \frac{1}{h}\int_{-\infty}^{\infty} dy\; y^2\; \mathrm{sinc}[\frac{1}{h}(y-y_{a'_y})]\; \mathrm{sinc}[\frac{1}{h}(y-y_{a_y})]\, \delta_{a'_x a_x} \right) \\ &= \frac{1}{2}\omega^2 (R_{a'_x a_x}\delta_{a'_y a_y} + R_{a'_y a_y}\delta_{a'_x a_x}). \end{aligned} \tag{4.2}$$

Since the integrals in the x and y dimensions are equal, it is enough to handle only one of them. Using the definition of the sinc functions, the integration in the x-dimension becomes

$$\begin{aligned} R_{a'a} &= \frac{1}{h}\int_{-M}^{M} dx\; x^2\, \frac{\sin[\frac{\pi}{h}(x-x_{a'})]}{\frac{\pi}{h}(x-x_{a'})}\, \frac{\sin[\frac{\pi}{h}(x-x_a)]}{\frac{\pi}{h}(x-x_a)} \\ &= \frac{h}{\pi^2}\int_{-M-x_{a'}}^{M-x_{a'}} dt\; (t+x_{a'})^2\, \frac{\sin(\frac{\pi}{h}t)\sin[\frac{\pi}{h}(t+x_{a'}-x_a)]}{t(t+x_{a'}-x_a)}, \end{aligned} \tag{4.3}$$

where the x-subscripts have been dropped for readability, and a change of variable $t = x - x_{a'}$ has been made. The integration limits have also been changed from $-\infty$ and ∞ to $-M$ and M, where $M \to \infty$, because the improper integrals will cause trouble ahead. Expanding $(t+x_{a'})^2$ yields three terms

$$\begin{aligned} R_{a'a} &= \frac{h}{\pi^2}\left(\int_{-M-x_{a'}}^{M-x_{a'}} dt\; t\, \frac{\sin(\frac{\pi}{h}t)\sin[\frac{\pi}{h}(t+x_{a'}-x_a)]}{t+x_{a'}-x_a} \right. \\ &\quad + 2x_{a'}\int_{-M-x_{a'}}^{M-x_{a'}} dt\, \frac{\sin(\frac{\pi}{h}t)\sin[\frac{\pi}{h}(t+x_{a'}-x_a)]}{t+x_{a'}-x_a} \\ &\quad \left. + x_{a'}^2\int_{-M-x_{a'}}^{M-x_{a'}} dt\, \frac{\sin(\frac{\pi}{h}t)\sin[\frac{\pi}{h}(t+x_{a'}-x_a)]}{t(t+x_{a'}-x_a)} \right) \\ &= \frac{h}{\pi^2}(I_1+I_2+I_3). \end{aligned}$$

Making a change of variable $s = t + x_{a'} - x_a$, it can be seen that the integrand in the second term, I_2, is odd, and since the integration is done in the limit $M \to \infty$, the integration results in $I_2 = 0$. The third term, I_3, can be easily calculated owing to the orthogonality of the Lagrange functions, and it gives

$$I_3 = \frac{\pi^2}{h}x_{a'}^2\, \delta_{a'a}. \tag{4.4}$$

The I_3 term clearly gives the diagonal contribution to the potential matrix, which is the part that the Gauss quadrature approximation takes into account.

The first term, I_1, is the one that causes all the trouble. The integrand in I_1 can be first simplified by noting that

$$\sin[\frac{\pi}{h}(t + x_{a'} - x_a)] = \sin[\frac{\pi}{h}t + \pi(a' - a)] = (-1)^{a'-a}\sin(\frac{\pi}{h}t),$$

where the definition (3.7) has been used, i.e. $x_a = ha$, where $a \in \mathbb{Z}$. Using this and making another change of variable, $s = t + x_{a'} - x_a$, leads to

$$\begin{aligned} I_1 &= (-1)^{a'-a}\int_{-M-x_{a'}}^{M-x_{a'}} dt\; t\,\frac{\sin^2(\frac{\pi}{h}t)}{t + x_{a'} - x_a} \\ &= (-1)^{a'-a}\int_{-M-x_a}^{M-x_a} ds\;(s - x_{a'} + x_a)\,\frac{\sin^2[\frac{\pi}{h}s + \pi(a' - a)]}{s} \\ &= (-1)^{a'-a}\left(\int_{-M-x_a}^{M-x_a} ds\;\sin^2(\frac{\pi}{h}s) \;+\; (x_a - x_{a'})\int_{-M-x_a}^{M-x_a} ds\;\frac{\sin^2(\frac{\pi}{h}s)}{s}\right). \end{aligned} \tag{4.5}$$

The second term goes to zero when $M \to \infty$ because the integrand is odd. Using the trigonometric identity $\sin^2(\frac{\pi}{h}s) = (1 - \cos(\frac{2\pi}{h}s))/2$, I_1 becomes

$$\begin{aligned} I_1 &= \frac{(-1)^{a'-a}}{2}\left(\int_{-M-x_a}^{M-x_a} ds - \int_{-M-x_a}^{M-x_a} ds\;\cos(\frac{2\pi}{h}s)\right) \\ &= (-1)^{a'-a}[M - \sin(\frac{2\pi}{h}M)] \\ &= (-1)^{a'-a}M. \end{aligned} \tag{4.6}$$

To get rid of the latter term, M is defined as $M = nh$, where $n \in \mathbb{Z}$ and $n \to \infty$. That term makes no difference anyway, because it is bounded by $-1 \leq \sin(\frac{2\pi}{h}M) \leq 1$.

Combining the terms I_1 (4.6) and I_3 (4.4) gives

$$R_{a'a} = x_{a'}^2\,\delta_{a'a} + (-1)^{a'-a}\frac{h}{\pi^2}M. \tag{4.7}$$

As $M \to \infty$, the potential matrix elements become infinitely large in magnitude. Therefore, it seems that the sinc basis would fail if the potential is parabolic. However, if the potential elements are approximated using the Gauss quadrature, the elements are finite, and the resulting single-particle states are reasonable. It will be shown next that the additional matrix with infinitely large elements alters the majority of the eigenvalues and eigenstates of the Hamiltonian matrix only slightly. The extremely regular form of the matrix is the reason for this.

4.2 Effect of Gauss quadrature in 1-D

4.2.1 Difference matrix and its eigendecomposition

For the sake of simplicity, the case of a one-dimensional Lagrange mesh is handled first and then the result is generalized to two dimensions. For a one-dimensional

mesh, the potential matrix elements are

$$\begin{aligned} V_{a'a} &= \frac{1}{2}\omega^2 R_{a'a} = \frac{1}{2}\omega^2 (x_{a'}^2\, \delta_{a'a} + (-1)^{a'-a}\frac{h}{\pi^2}M) \\ &= V(x_a)\delta_{a'a} + (-1)^{a'-a}\frac{h\omega^2}{2\pi^2}M. \end{aligned} \tag{4.8}$$

The differences to the elements that are obtained using the Gauss quadrature are given by

$$\Delta_{a'a} = (-1)^{a'-a}\frac{h\omega^2}{2\pi^2}M. \tag{4.9}$$

Using these differences, the Hamiltonian with analytically calculated potential matrix elements can be written as

$$\mathrm{H} = \mathrm{H}_0 + \Delta, \tag{4.10}$$

where H_0 is the Hamiltonian that contains only the diagonal part of the potential matrix, i.e. the part that the Gauss quadrature approximation gives. Therefore, the effect of the Gauss quadrature approximation on the eigenvalues and eigenvectors of the Hamiltonian is given by the difference between those of H_0 and H. Since the elements of Δ are infinitely large in magnitude, one would guess that the eigenvalues of H would also be infinitely large, or at least the difference between the eigenvalues of H_0 and H would be enormous. Fortunately, this is not the case.

Generally, the eigenvalues of a sum of matrices are difficult to estimate based on the eigenvalues of the separate matrices. However, some boundaries for the eigenvalues of the sum of two matrices can be obtained using the theorems explained by Wilkinson [31]. The regular form of the difference matrix Δ,

$$\Delta = \frac{h\omega^2}{2\pi^2}M \begin{bmatrix} 1 & -1 & 1 & \cdots & (-1)^{N-1} \\ -1 & 1 & -1 & \cdots & (-1)^{N} \\ 1 & -1 & 1 & \cdots & (-1)^{N-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ (-1)^{N-1} & (-1)^{N} & (-1)^{N-1} & \cdots & 1 \end{bmatrix}, \tag{4.11}$$

is the key to the calculation of the boundaries. In addition to being symmetric, this matrix has three distinctive properties:

1. Δ is a rank-1 matrix,

2. Δ is a circulant matrix, if N is even,

3. Δ is a skew right circulant (or -1-factor circulant) matrix, if N is odd [32].

In the one dimensional case, the rank-1 property is enough, but it should be noted that the circulant property becomes handy in the two dimensional case. Since Δ is a rank-1 matrix, i.e. it has only one linearly independent column, it has exactly

one non-zero eigenvalue. This fact is stated by the rank-nullity theorem, which says that for a $n \times n$ matrix A [27]

$$\dim(\text{null(A)}) + \text{rank(A)} = n,$$

where null(A) denotes the null space of A. Null space is defined as

$$\text{null(A)} = \{\mathbf{y} \in \mathbb{R}^n \,:\, \text{A}\mathbf{y} = 0\},$$

i.e for any vector $\mathbf{y}$ in the null space, $\text{A}\mathbf{y} = 0$. Since Δ is an $N \times N$ matrix and $\text{rank}(\Delta) = 1$, $\dim(\text{null}(\Delta)) = N - 1$. Therefore, there must be $N - 1$ orthogonal vectors $\mathbf{y}_i$ that span the null space of Δ, i.e.

$$\Delta \mathbf{y}_i = 0, \quad i \in \{1, 2, \ldots, N-1\}.$$

These vectors are eigenvectors of Δ with a zero eigenvalue. The vector $\mathbf{x} \in \mathbb{R}^N$ that is orthogonal to all $\mathbf{y}_i$ must be the only eigenvector of Δ which has a non-zero eigenvalue. As the matrix is also symmetric, it can be factorized to the form

$$\Delta = \mathbf{x}\, \sigma\, \mathbf{x}^{\text{T}}, \tag{4.12}$$

where σ is the only non-zero eigenvalue and $\mathbf{x}$ is the corresponding normalized eigenvector. The eigenvector must be parallel to the only linearly independent column and, properly normalized, it can be written as

$$\mathbf{x} = \frac{1}{\sqrt{N}}[1 \quad -1 \quad 1 \quad \cdots \quad (-1)^{N-1}]^{\text{T}}. \tag{4.13}$$

It can be easily seen that the corresponding eigenvalue must be

$$\sigma = \frac{h\omega^2}{2\pi^2} MN. \tag{4.14}$$

Since the other eigenvalues are zero, $N - 1$ mutually orthonormal vectors can be chosen in addition to $\mathbf{x}$ to form a basis in which Δ is diagonal, i.e.

$$\Delta = \text{U}\Sigma\text{U}^{\text{T}}, \tag{4.15}$$

where $\Sigma = \text{diag}(\sigma, 0, 0, \ldots)$ and the first column of U is $\mathbf{x}$. The complete Hamiltonian in (4.10) can be transformed into the basis of U to obtain

$$\text{H} = \text{UH}'\text{U}^{\text{T}}, \tag{4.16}$$

where

$$\text{H}' = \text{U}^{\text{T}}(\text{H}_0 + \Delta)\text{U} = \text{H}_0' + \Sigma. \tag{4.17}$$

Since H and H′ are similar, they share the same eigenvalues. In this basis, the difference between the Hamiltonians, H′ and H_0', is only in the first diagonal element because

$$\text{H}' = \text{H}_0' + \begin{bmatrix} \sigma & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}. \tag{4.18}$$

4.2.2 Perturbation of eigenvalues with infinite diagonal element

The Wilkinson's treatment of a perturbation of eigenvalues by a symmetric rank-1 matrix [31] also proceeds through the form where the perturbation matrix has only one diagonal element, and after a few further steps it results in

$$\lambda_i(\mathrm{H}') \in [\lambda_i(\mathrm{H}'_0), \lambda_{i-1}(\mathrm{H}'_0)], \quad i \in [2, N] \tag{4.19}$$

and

$$\begin{cases} \lambda_i(\mathrm{H}') = \lambda_i(\mathrm{H}'_0) + m_i\sigma, & i \in [1, N] \\ \sum_{i=1}^{N} m_i = 1 \end{cases} \tag{4.20}$$

where $\lambda_i(\mathrm{H}')$ denotes the ith eigenvalue of H'. The eigenvalues are sorted in decreasing order. From the bounds of Eq. (4.19), it is evident that the eigenvalues from $\lambda_2(\mathrm{H}'_0)$ to $\lambda_N(\mathrm{H}'_0)$ can grow only as large as the next largest eigenvalue of H'_0 when the rank-1 matrix is added. $\lambda_1(\mathrm{H}')$ is the only eigenvalue that is not bounded by the eigenvalues of H'_0. Since the eigenvalues of H' are determined by Eq. (4.20) and the eigenvalues from $\lambda_2(\mathrm{H}')$ to $\lambda_N(\mathrm{H}')$ must be finite, the coefficients m_i must be chosen so that

$$m_i \sim \frac{1}{\sigma}, \quad i \in [2, N]. \tag{4.21}$$

This, together with the sum condition for the m_i, implies that $m_1 = 1$, because $\sigma \to \infty$ when $M \to \infty$. Therefore, the largest eigenvalue of H' approaches infinity, whereas the others stay finite.

The previous result for the eigenvalues of H' only states that $N - 1$ of the eigenvalues are bounded by the eigenvalues of H'_0, whereas the largest one approaches infinity. Since the gap between two consecutive eigenvalues, $\lambda_i(\mathrm{H}'_0)$ and $\lambda_{i-1}(\mathrm{H}'_0)$, can be very large, this result is obviously not good enough. For the Gauss quadrature approximation to be reasonable, the difference between the eigenvalues of H' and H'_0, omitting the infinite eigenvalue of H', should be small and it should converge to zero in the limit $L \to \infty$ and $h \to 0$. A qualitative proof will be given next to convince that this indeed is true.

Adding the one infinitely large diagonal element σ to the Hamiltonian H'_0 causes the first basis vector of the Hamiltonian matrix to become one of the eigenvectors of the matrix. This is quite obvious since the off-diagonal elements in the first column of the Hamiltonian are negligibly small compared to the infinite diagonal element and thus they can be ignored. The first basis vector of the original Hamiltonian H is the first column of U, which is $\mathbf{x}$, the eigenvector corresponding to the nonzero eigenvalue of Δ. Consequently, the eigenvector corresponding to the infinite eigenvalue of H is $\mathbf{x}$, and the other eigenvectors must be orthogonal to it. The eigenvectors of H_0 are thus transformed due to the infinite diagonal element so that their component to the direction of $\mathbf{x}$ is removed. It is safe to assume that the eigenvectors which are almost orthogonal to the vector $\mathbf{x}$ to begin with do not change notably.

The following analysis shows that especially the eigenvectors that correspond to the lowest eigenvalues of H_0 are approximately orthogonal to $\mathbf{x}$ and, therefore,

they stay almost unchanged when the Δ matrix is added. The measure for the orthogonality of the eigenvectors of H_0 and the vector $\mathbf{x}$ is given by the inner product $\mathbf{v}_j^{\mathrm{T}}\mathbf{x}$, where $\mathbf{v}_j$ is the jth eigenvector of H_0. In Appendix A it is shown that the magnitude of the inner product is bounded by

$$\begin{aligned}|\mathbf{v}_j^{\mathrm{T}}\mathbf{x}| \leq \; & \frac{1}{\sqrt{N}}\left[h^{1/2}\frac{1}{2}|\Psi_j(t_{N+1}) - \Psi_j(t_1)| + h^{3/2}\frac{3}{4}|\Psi_j'(t_{N+1}) - \Psi_j'(t_1)| + \right. \\ & + h^{5/2}\left(\frac{L}{4}\max_{\xi\in[t_1,t_{N+1}]}|\Psi_j'''(\xi)| + \frac{1}{4}|\Psi_j''(t_{N+1}) - \Psi_j''(t_1)|\right) + \\ & \left. + h^{7/2}\frac{L}{12}\max_{\xi\in[t_1,t_{N+1}]}|\Psi_j^{(4)}(\xi)|\right],\end{aligned} \tag{4.22}$$

where t_1 is the first mesh point and t_{N+1} the mesh point next to the last one. As $h \to 0$, each term approaches zero, the first one being the slowest. Another argument can be applied to prove that the terms with values of wave function or its derivatives at the mesh points t_1 and t_{N+1} should be negligible. The values of the wave function at the edges of the simulation area, t_1 and t_{N+1}, should be approximately zero because, otherwise, the simulation area is not large enough and the basis functions cannot represent the wave function properly. Since the Lagrange functions are infinitely differentiable, the wave function is a smooth function, and therefore, the derivatives of the wave function should also be approximately zero at the edges of the simulation area. Although the derived upper bound for the inner product $\mathbf{v}_j^{\mathrm{T}}\mathbf{x}$ is rather crude, it clearly shows that the inner product goes to zero as $h \to 0$ and it is small even for $h \approx 1$.

The previous analysis shows that the magnitude of the inner product $\mathbf{v}_j^{\mathrm{T}}\mathbf{x}$ can be made arbitrarily small by reducing the mesh spacing. Increasing the size of the simulation area also makes the inner product smaller because the values of the wave function and its derivatives at the outermost mesh points decrease. As the wave functions of the higher electron states tend to oscillate more and extend further than the lower ones, they may be nonzero at the edges of the simulation area and the maximum values of the derivatives are higher. Therefore, the inner product $\mathbf{v}_j^{\mathrm{T}}\mathbf{x}$ is typically small for a number of lowest states but significant for the higher ones. The interpretation of this result is that the vector $\mathbf{x}$, the eigenvector of the Δ matrix, is almost orthogonal to the eigenvectors corresponding to a certain number of lowest eigenvalues of the Hamiltonian H_0. As was reasoned before, these eigenvectors stay almost constant when the Δ matrix is added to the Hamiltonian H_0. As the change in the eigenvectors is negligible, one would guess that the eigenvalues are not altered significantly either. The change in the eigenvalues is studied numerically in the next section.

4.2.3 Numerical study

The eigenvalues of the Hamiltonians with and without the Gauss quadrature approximation, H_0 and H respectively, were calculated with Mathematica computer algebra system. The Hamiltonian matrices were constructed for a 1-D parabolic potential which was in the center of the simulation area. The length of the simulation

area was chosen to be sufficient for a couple of first eigenstates and kept constant, and the number of basis functions was varied. When the confinement strength of the potential was chosen to be $\hbar\omega_0 = 3$ meV, a simulation area length of $L = 200$ nm was found to be adequate. The eigenvalues of H were calculated analytically and the limit $M \to \infty$ was taken in the end.

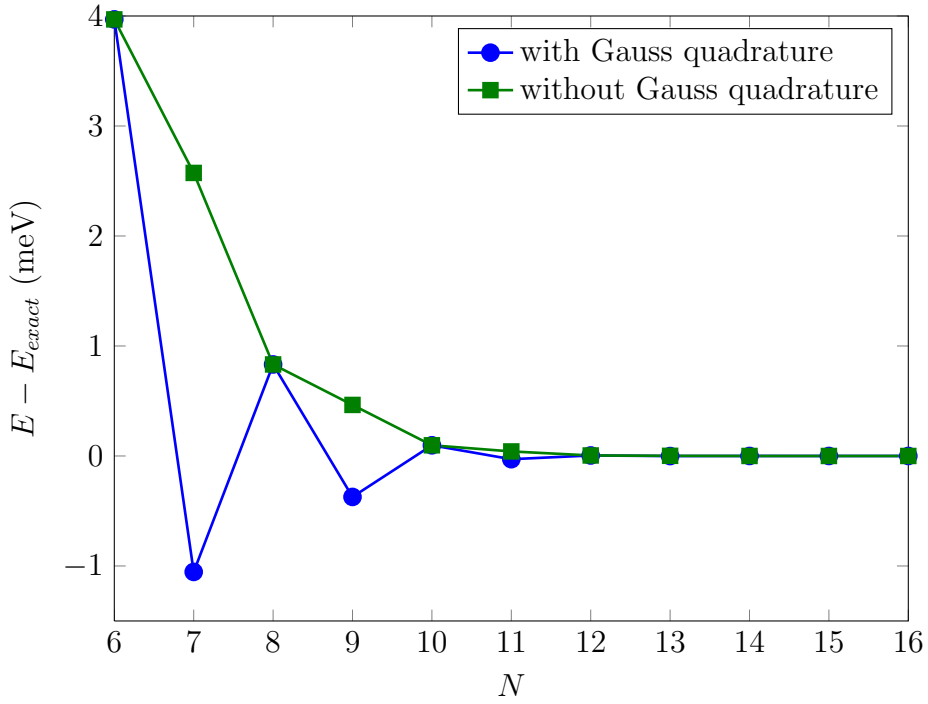

Figure 4.1: The convergence of the third eigenvalue of H_0 and H towards the correct energy of the third state as a function of the number of basis states, N. The confinement strength was chosen to be $\hbar\omega_0 = 3$ meV and the simulation area length $L = 200$ nm. The correct energy of the third state of a quantum harmonic oscillator is $E_{exact} = 7.5$ meV.

The exact energies of a particle in a parabolic potential, i.e. the quantum harmonic oscillator, are known. Fig. 4.1 shows the convergence of the third eigenvalue of H_0 and H towards the correct energy of the third state, which is $E_{exact} = 7.5$ meV. The convergence is clearly different depending on whether Gauss quadrature approximation is used or not. The eigenvalue which is calculated without the quadrature approximation is always above the correct one, whereas the other one goes below the correct eigenenergy. Therefore, the Gauss quadrature approximation clearly breaks up the variational principle, which states that the energy of each trial wave function is above the energy of the correct one.

Fig. 4.2 shows the absolute differences between eigenvalues calculated with and without the Gauss quadrature approximation. The differences were calculated for the four lowest eigenvalues. When N is even, it makes no difference which way the first and the third eigenvalues are calculated as $\Delta E_1 \approx 0$ and $\Delta E_3 \approx 0$ up to numerical precision. When N is even, ΔE_1 and ΔE_3 decrease exponentially fast as

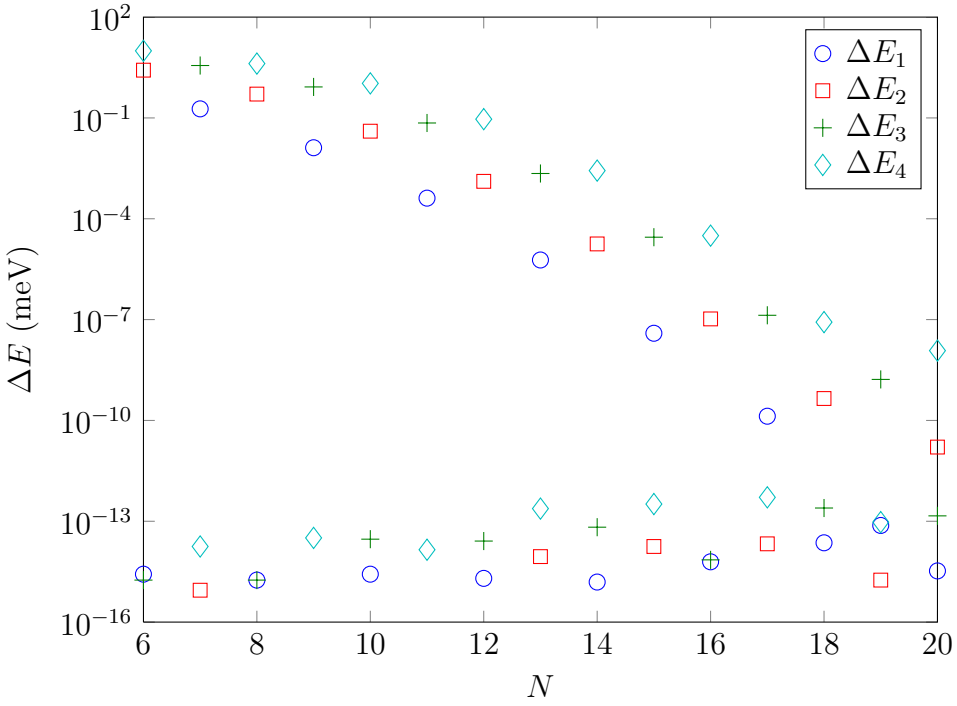

Figure 4.2: The absolute difference between four eigenvalue pairs calculated with and without the Gauss quadrature approximation. $\Delta E_i = |E_i^{\mathrm{H}} - E_i^{\mathrm{H}_0}|$.

a function of N. ΔE_2 and ΔE_4 behave the same way except that the parity of N is the other way around. Both dependences on N can be related to the magnitude of the inner product $\mathbf{v}_j^{\mathrm{T}}\mathbf{x}$.

The dependence on the parity of N can be explained by writing the inner product as

$$\mathbf{v}_j^{\mathrm{T}}\mathbf{x} = \frac{1}{\sqrt{N}}\sqrt{h}\sum_{i=1}^{N}(-1)^{i-1}\Psi_j(x_{-\frac{N+1}{2}+i}), \tag{4.23}$$

which can be derived by following the beginning of Appendix A. The discretized position variable $x_{-\frac{N+1}{2}+i}$ can be also written as

$$x_{-\frac{N+1}{2}+i} = -\frac{L}{2} - \frac{h}{2} + hi.$$

When N is even, the sum can be divided into two parts and the sum indices can be

conveniently manipulated to derive

$$\begin{aligned}
\mathbf{v}_j^{\mathrm{T}}\mathbf{x} &= \frac{1}{\sqrt{N}}\sqrt{h}\left[\sum_{i=1}^{N/2}(-1)^{i-1}\Psi_j(x_{-\frac{N+1}{2}+i}) + \sum_{i=N/2+1}^{N}(-1)^{i-1}\Psi_j(x_{-\frac{N+1}{2}+i})\right] \\
&= \frac{1}{\sqrt{N}}\sqrt{h}\left[\sum_{i=1}^{N/2}(-1)^{i}\Psi_j(x_{\frac{1}{2}-i}) + \sum_{i=1}^{N/2}(-1)^{i-1}\Psi_j(x_{-\frac{1}{2}+i})\right] \\
&= \frac{1}{\sqrt{N}}\sqrt{h}\sum_{i=1}^{N/2}(-1)^{i}\left[\Psi_j(-x_{-\frac{1}{2}+i}) - \Psi_j(x_{-\frac{1}{2}+i})\right].
\end{aligned} \tag{4.24}$$

Since the potential is symmetric and centred at the origin, the wave functions must be either symmetric or antisymmetric. In particular, the first and the third eigenstates of the quantum harmonic oscillator are symmetric and the second and the fourth eigenstates are antisymmetric. For a symmetric wave function, $\Psi_j(-x) = \Psi_j(x)$. Therefore, $\mathbf{v}_j^{\mathrm{T}}\mathbf{x} = 0$ for symmetric wave functions when N is even. Similar arguments can be applied to show that $\mathbf{v}_j^{\mathrm{T}}\mathbf{x} = 0$ for antisymmetric wave functions when N is odd. This shows that there clearly is a connection between the magnitude of the inner product $\mathbf{v}_j^{\mathrm{T}}\mathbf{x}$ and the difference between the eigenvalues that are calculated with and without the Gauss quadrature approximation. When the inner product is zero, it makes no difference whether the quadrature approximation is used or not.

If the points where $\Delta E_i \approx 0$ are ignored, the values of ΔE_i decrease exponentially fast as a function of N. In the last section, a bound for the magnitude of $\mathbf{v}_j^{\mathrm{T}}\mathbf{x}$ was introduced in Eq. (4.22). The bound shows that $|\mathbf{v}_j^{\mathrm{T}}\mathbf{x}|$ decreases when the mesh spacing h decreases. Since $h = L/N$, $|\mathbf{v}_j^{\mathrm{T}}\mathbf{x}|$ decreases as a function of N. Therefore, the magnitude of the inner product and the values of ΔE_i are clearly correlated and the numerical results agree with the reasoning of the previous section.

4.3 Generalization to 2-D

Now that the error due to Gauss quadrature approximation in 1-D has been shown to vanish when $h \to 0$ and to be small for a number of lowest eigenstates for reasonably small h, the result is generalized to 2-D. The analytically calculated potential matrix elements for a two dimensional mesh are given by

$$\begin{aligned}
V_{\mathbf{a'a}} &= \frac{1}{2}\omega^2(R_{a'_x a_x}\delta_{a'_y a_y} + R_{a'_y a_y}\delta_{a'_x a_x}) \\
&= \frac{1}{2}\omega^2\left[(x_{a'_x} + x_{a'_y})\delta_{a'_x a_x}\delta_{a'_y a_y} + \frac{h}{\pi^2}M\left((-1)^{a'_x - a_x}\delta_{a'_y a_y} + (-1)^{a'_y - a_y}\delta_{a'_x a_x}\right)\right] \\
&= V(\mathbf{r_a})\delta_{a'_x a_x}\delta_{a'_y a_y} + \frac{h\omega^2}{2\pi^2}M\left((-1)^{a'_x - a_x}\delta_{a'_y a_y} + (-1)^{a'_y - a_y}\delta_{a'_x a_x}\right),
\end{aligned} \tag{4.25}$$

which is obtained by combining Eq. (4.2) with Eq. (4.7). The differences to the elements calculated using the Gauss quadrature approximation are given by

$$\Delta_{a'_x a'_y a_x a_y} = \begin{cases} \frac{h\omega^2}{\pi^2} M & , a'_x = a_x, a'_y = a_y \\ (-1)^{a'_y - a_y} \frac{h\omega^2}{2\pi^2} M & , a'_x = a_x, a'_y \neq a_y \\ (-1)^{a'_x - a_x} \frac{h\omega^2}{2\pi^2} M & , a'_x \neq a_x, a'_y = a_y \\ 0 & , a'_x \neq a_x, a'_y \neq a_y \end{cases}. \tag{4.26}$$

Therefore, the Hamiltonian matrix with the analytically calculated potential matrix elements can be written as

$$\mathrm{H} = \mathrm{H}_0 + \Delta. \tag{4.27}$$

For the following analysis, the Δ matrix is written as

$$\Delta = \frac{h\omega^2}{2\pi^2} M\mathcal{B}, \tag{4.28}$$

where $\mathcal{B}$ is a block matrix

$$\mathcal{B} = \begin{bmatrix} \mathrm{C} & -\mathrm{I} & \mathrm{I} & \cdots & (-1)^{N-1}\mathrm{I} \\ -\mathrm{I} & \mathrm{C} & -\mathrm{I} & \cdots & (-1)^{N}\mathrm{I} \\ \mathrm{I} & -\mathrm{I} & \mathrm{C} & \cdots & (-1)^{N-1}\mathrm{I} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ (-1)^{N-1}\mathrm{I} & (-1)^{N}\mathrm{I} & (-1)^{N-1}\mathrm{I} & \cdots & \mathrm{C} \end{bmatrix}. \tag{4.29}$$

The C blocks are $N \times N$ matrices

$$\mathrm{C} = \begin{bmatrix} 2 & -1 & 1 & \cdots & (-1)^{N-1} \\ -1 & 2 & -1 & \cdots & (-1)^{N} \\ 1 & -1 & 2 & \cdots & (-1)^{N-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ (-1)^{N-1} & (-1)^{N} & (-1)^{N-1} & \cdots & 2 \end{bmatrix}, \tag{4.30}$$

and I are $N \times N$ identity matrices.

The additional matrix Δ in the 1-D case was a rank-1 matrix for which the only non-zero eigenvalue and the corresponding eigenvector were easy to obtain. In the 2-D case, one must calculate the non-zero eigenvalues and the eigenvectors of the block matrix $\mathcal{B}$. Fortunately, $\mathcal{B}$ is a factor block circulant matrix and the diagonalization of it is relatively simple and can be done analytically.

4.3.1 Diagonalization of factor block circulant matrix $\mathcal{B}$

A square matrix is called circulant if each subsequent row is obtained from the previous one by a cyclic shift. A factor circulant matrix is a circulant matrix where each element below the diagonal is multiplied by a common factor. An A-factor

block circulant matrix of size $mn \times mn$ is defined as [33]

$$\mathrm{circ_A}(\mathrm{C}_1, \mathrm{C}_2, \ldots, \mathrm{C}_m) \equiv \begin{bmatrix} \mathrm{C}_1 & \mathrm{C}_2 & \mathrm{C}_3 & \cdots & \mathrm{C}_m \\ \mathrm{AC}_m & \mathrm{C}_1 & \mathrm{C}_2 & \cdots & \mathrm{C}_{m-1} \\ \mathrm{AC}_{m-1} & \mathrm{AC}_m & \mathrm{C}_1 & \cdots & \mathrm{C}_{m-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathrm{AC}_2 & \mathrm{AC}_3 & \mathrm{AC}_4 & \cdots & \mathrm{C}_1 \end{bmatrix}, \tag{4.31}$$

where the factor A is a $n \times n$ matrix. The blocks C_i can in general be any $n \times n$ matrices, but for the current purposes they are assumed to be a-factor circulant matrices defined as

$$\mathrm{C}_i = \mathrm{circ}_a(c_1, c_2, \ldots, c_m) \equiv \begin{bmatrix} c_1 & c_2 & c_3 & \cdots & c_n \\ ac_n & c_1 & c_2 & \cdots & c_{n-1} \\ ac_{n-1} & ac_n & c_1 & \cdots & c_{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ ac_2 & ac_3 & ac_4 & \cdots & c_1 \end{bmatrix}. \tag{4.32}$$

The matrix $\mathcal{B}$ is a block circulant, i.e. I-factor block circulant, matrix if N is even,

$$\mathcal{B}_\mathrm{I} = \mathrm{circ_I}(\mathrm{C}, -\mathrm{I}, \mathrm{I}, \ldots, -\mathrm{I}), \tag{4.33}$$

and a $-$I-factor block circulant matrix if N is odd,

$$\mathcal{B}_{-\mathrm{I}} = \mathrm{circ_{-I}}(\mathrm{C}, -\mathrm{I}, \mathrm{I}, \ldots, \mathrm{I}). \tag{4.34}$$

The C matrix is also a circulant or a -1-factor circulant matrix depending on the parity of N, and the identity matrix can be interpreted either as a circulant or a -1-factor circulant matrix. Therefore, the blocks of $\mathcal{B}$ are also circulant if N is even and -1-factor circulant if N is odd.

Claeyssen and dos Santos Leal [33] have derived the diagonalized form for a general A-factor block circulant matrix. Since the factor matrix of $\mathcal{B}$ is either I or $-$I, the most general result is not needed. Therefore, a simpler result for I and $-$I is derived by following the steps of the article.

The block Vandermonde matrix of the $n \times n$ matrices H_k is defined as

$$\mathcal{V}_n(\mathrm{H}_1, \mathrm{H}_2, \ldots, \mathrm{H}_m) \equiv \begin{bmatrix} \mathrm{I} & \mathrm{I} & \mathrm{I} & \cdots & \mathrm{I} \\ \mathrm{H}_1 & \mathrm{H}_2 & \mathrm{H}_3 & \cdots & \mathrm{H}_m \\ \mathrm{H}_1^2 & \mathrm{H}_2^2 & \mathrm{H}_3^2 & \cdots & \mathrm{H}_m^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathrm{H}_1^{m-1} & \mathrm{H}_2^{m-1} & \mathrm{H}_3^{m-1} & \cdots & \mathrm{H}_m^{m-1} \end{bmatrix}. \tag{4.35}$$

Let

$$\omega = e^{2\pi i/m}$$

denote the basic mth root of unity and K the principal mth root of the factor matrix A. K is thus defined as [34]

$$\mathrm{K} \in \mathbb{C}^{n \times n}, \quad \mathrm{K}^m = A, \quad \arg[\sigma(\mathrm{K})] \in [-\pi/m, \pi/m],$$

where $\sigma(\mathrm{K}) = \lambda_i, i = 1, \ldots, n$ is the eigenspectrum of K. Principal root of a matrix is thus the root for which the complex arguments of the eigenvalues are as small as possible.

The most important basic result that Clayessen and dos Santos Leal proved is that the A-factor block circulant matrix $\mathcal{C} = \mathrm{circ_A}(\mathrm{C}_1, \mathrm{C}_2, \ldots, C_m)$ can be block diagonalized as

$$\mathcal{C} = \mathcal{V}_\mathrm{A} \mathcal{P}(\mathcal{D}_\mathrm{A}) \mathcal{V}_\mathrm{A}^{-1}, \tag{4.36}$$

where

$$\mathcal{V}_\mathrm{A} = \mathcal{V}_n(\mathrm{K}, \omega\mathrm{K}, \ldots, \omega^{m-1}\mathrm{K}), \tag{4.37}$$

$$\mathcal{D}_\mathrm{A} = \mathrm{diag}(\mathrm{K}, \omega\mathrm{K}, \ldots, \omega^{m-1}\mathrm{K}) \tag{4.38}$$

and

$$\mathcal{P}(\mathcal{D}_\mathrm{A}) = \sum_{k=0}^{m-1} \mathrm{C}_{k+1} \mathcal{D}_\mathrm{A}^k. \tag{4.39}$$

The product of an $n \times n$ matrix C_{k+1} with an $mn \times mn$ block matrix $\mathcal{D}_\mathrm{A}^k$ is understood as a block-wise operation.

The following properties of block diagonal matrices are needed in the subsequent manipulations. Let $\mathcal{X} = \mathrm{diag}(\mathrm{X}_1, \mathrm{X}_2, \ldots, \mathrm{X}_m)$ and $\mathcal{Y} = \mathrm{diag}(\mathrm{Y}_1, \mathrm{Y}_2, \ldots, \mathrm{Y}_m)$ be $mn \times mn$ block diagonal matrices, where the blocks are $n \times n$ matrices. The product of these two block diagonal matrices is

$$\mathcal{X}\mathcal{Y} = \mathrm{diag}(\mathrm{X}_1\mathrm{Y}_1, \mathrm{X}_2\mathrm{Y}_2, \ldots, \mathrm{X}_m\mathrm{Y}_m),$$

i.e. the multiplication is done block by block. This can be verified by expanding the blocks X_i and Y_i and performing a direct multiplication of the $\mathcal{X}$ and $\mathcal{Y}$ matrices. From this block-wise multiplication, it follows that

$$\mathcal{X}^k = \mathrm{diag}(\mathrm{X}_1^k, \mathrm{X}_2^k, \ldots, \mathrm{X}_m^k).$$

Also

$$\begin{aligned} \mathcal{X}\ \mathrm{diag}(\mathrm{X}_1^{-1}, \mathrm{X}_2^{-1}, \ldots, \mathrm{X}_m^{-1}) &= \mathrm{diag}(\mathrm{X}_1\mathrm{X}_1^{-1}, \mathrm{X}_2\mathrm{X}_2^{-1}, \ldots, \mathrm{X}_m\mathrm{X}_m^{-1}) \\ &= \mathrm{diag}(\mathrm{I}, \mathrm{I}, \ldots, \mathrm{I}) = \mathcal{I}, \end{aligned}$$

which indicates that the inverse matrix of $\mathcal{X}$ is

$$\mathcal{X}^{-1} = \mathrm{diag}(\mathrm{X}_1^{-1}, \mathrm{X}_2^{-1}, \ldots, \mathrm{X}_m^{-1})$$

Due to these properties of block diagonal matrices

$$\mathcal{D}_\mathrm{A}^k = \mathrm{diag}(\mathrm{K}^k, (\omega\mathrm{K})^k, \ldots, (\omega^{m-1}\mathrm{K})^k),$$

and $\mathcal{P}(\mathcal{D}_\mathrm{A})$ can be written as

$$\mathcal{P}(\mathcal{D}_\mathrm{A}) = \mathrm{diag}(\mathcal{P}(\mathrm{K}), \mathcal{P}(\omega\mathrm{K}), \ldots, \mathcal{P}(\omega^{m-1}\mathrm{K})) = \mathrm{diag}(\mathrm{M}_1, \mathrm{M}_2, \ldots, \mathrm{M}_m). \tag{4.40}$$

If $\mathcal{X}$ is a block diagonal matrix as previously, it is easy to see that the following equivalence holds

$$\mathcal{X}^{-1}\mathcal{P}(\mathcal{D}_{\mathrm{A}})\mathcal{X} = \mathrm{diag}(\mathrm{X}_1^{-1}\mathrm{M}_1\mathrm{X}_1,\, \mathrm{X}_2^{-1}\mathrm{M}_2\mathrm{X}_2,\, \ldots,\, \mathrm{X}_m^{-1}\mathrm{M}_m\mathrm{X}_m).$$

This means that each block M_i in the block diagonal matrix $\mathcal{P}(\mathcal{D}_{\mathrm{A}})$ can be diagonalized individually by choosing the blocks X_i in the block diagonal matrix $\mathcal{X}$ so that

$$\mathrm{X}_i^{-1}\mathrm{M}_i\mathrm{X}_i = \Lambda_i,$$

where $\Lambda_i = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_m)$ and λ_i are the eigenvalues of M_i.

In our case, $\mathcal{B}$ is an $NN \times NN$ block matrix that is I-factor circulant when N is even and $-$I-factor circulant when N is odd. The principal Nth roots of I and $-$I are $\mathrm{K_I} = e^{0\pi i/N}\mathrm{I} = \mathrm{I}$ and $\mathrm{K}_{-\mathrm{I}} = e^{\pi i/N}\mathrm{I} = \omega^{1/2}\mathrm{I}$.

Let us assume that N is even so that $\mathcal{B}$ is a I-factor circulant matrix. The block matrix $\mathcal{B}$ with an odd N, i.e. a $-$I-factor block circulant matrix, can be diagonalized using the principal Nth root $\mathrm{K}_{-\mathrm{I}} = \omega^{1/2}\mathrm{I}$ instead of $\mathrm{K_I} = \mathrm{I}$. The diagonalization of the $-$I-factor block circulant matrix is not presented in this thesis, but the eigenvalues and eigenvectors turn out to have the same form as those of the I-factor block circulant matrix. The blocks of the block diagonal matrix $\mathcal{P}(\mathcal{D}_{\mathrm{I}})$ are

$$\mathrm{M_i} = \mathcal{P}(\omega^{i-1}\mathrm{I}) = \sum_{k=0}^{N-1} \mathrm{C}_{k+1}(\omega^{i-1}\mathrm{I})^k = \sum_{k=0}^{N-1} \mathrm{C}_{k+1}\omega^{(i-1)k}. \tag{4.41}$$

Since the matrices C_k are circulant, they are diagonalized similarly as the block circulant matrix, i.e.

$$\mathrm{C}_k = \mathrm{V}_a\mathcal{P}_k(\mathrm{D}_a)\mathrm{V}_a^{-1}, \tag{4.42}$$

where

$$\mathrm{V}_a = \mathrm{V}(\sqrt[N]{a}, \omega\sqrt[N]{a}, \ldots, \omega^{N-1}\sqrt[N]{a}), \tag{4.43}$$

$$\mathrm{D}_a = \mathrm{diag}(\sqrt[N]{a}, \omega\sqrt[N]{a}, \ldots, \omega^{N-1}\sqrt[N]{a}) \tag{4.44}$$

and

$$\mathcal{P}_k(\mathrm{D}_a) = \sum_{l=0}^{N-1} c_{l+1}^k \mathrm{D}_a^l. \tag{4.45}$$

c_l^k are the elements of the circulant matrix C_k, i.e.

$$\mathrm{C}_k = \mathrm{circ}(c_1^k, c_2^k, \ldots, c_N^k).$$

$\mathrm{V}(h_1, h_2, \ldots, h_N)$ denotes the basic, block size 1, Vandermonde matrix. Since the matrices C_k are basic circulant matrices, the factor a is 1 and $\sqrt[N]{a} = 1$. The matrices M_i can now be diagonalized by diagonalizing the circulant matrices C_k in the sum

$$\begin{aligned} \mathrm{M}_i &= \sum_{k=0}^{N-1} \mathrm{C}_{k+1}\omega^{(i-1)k} = \sum_{k=0}^{N-1} \mathrm{V}_1\mathcal{P}_{k+1}(\mathrm{D}_1)\mathrm{V}_1^{-1}\omega^{(i-1)k} \\ &= \mathrm{V}_1\left(\sum_{k=0}^{N-1}\sum_{l=0}^{N-1} c_{l+1}^{k+1}\mathrm{D}_1^l\omega^{(i-1)k}\right)\mathrm{V}_1^{-1} = \mathrm{V}_1\Lambda_i\mathrm{V}_1^{-1}, \end{aligned} \tag{4.46}$$

where

$$\Lambda_i = \mathrm{diag}(\lambda_1^i, \lambda_2^i, \ldots, \lambda_N^i)$$

and

$$\lambda_j^i = \sum_{k=0}^{N-1}\sum_{l=0}^{N-1} c_{l+1}^{k+1}\omega^{(j-1)l}\omega^{(i-1)k}. \tag{4.47}$$

The diagonalized form of $\mathcal{P}(\mathcal{D}_\mathrm{I})$ is now

$$\mathcal{P}(\mathcal{D}_\mathrm{I}) = \mathcal{V}_1\Lambda\mathcal{V}_1^{-1}, \tag{4.48}$$

where

$$\mathcal{V}_1 = \mathrm{diag}(\mathrm{V}_1, \mathrm{V}_1, \ldots, \mathrm{V}_1)$$

and

$$\Lambda = \mathrm{diag}(\Lambda_1, \Lambda_2, \ldots, \Lambda_N).$$

The complete diagonalization of the I-factor block circulant matrix $\mathcal{B}$ is thus

$$\mathcal{B} = \mathcal{V}_\mathrm{I}\mathcal{P}(\mathcal{D}_\mathrm{I})\mathcal{V}_\mathrm{I}^{-1} = \mathcal{V}_\mathrm{I}\mathcal{V}_1\Lambda\mathcal{V}_1^{-1}\mathcal{V}_\mathrm{I}^{-1} = \mathcal{V}_\mathrm{I}\mathcal{V}_1\Lambda(\mathcal{V}_\mathrm{I}\mathcal{V}_1)^{-1}. \tag{4.49}$$

4.3.2 Explicit eigenvalues

The expression for the eigenvalues can be simplified by taking into account the form of the circulant blocks of $\mathcal{B}$,

$$\mathrm{C}_1 = \mathrm{circ}(2, -1, 1, \ldots, -1)$$

and

$$\mathrm{C}_k = \begin{cases} \mathrm{I} = \mathrm{circ}(1, 0, \ldots, 0) & , \text{ if } k \text{ is odd and } k \neq 1 \\ -\mathrm{I} = \mathrm{circ}(-1, 0, \ldots, 0) & , \text{ if } k \text{ is even} \end{cases}.$$

Therefore, the components of the circulant blocks depend on the indices k and l so that

$$c_{l+1}^{k+1} = \begin{cases} 0 & , \quad k > 0,\, l > 0 \\ 2 & , \quad k = 0,\, l = 0 \\ (-1)^l & , \quad k = 0,\, l > 0 \\ (-1)^k & , \quad k > 0,\, l = 0 \end{cases}. \tag{4.50}$$

The eigenvalues are thus given by

$$\begin{aligned} \lambda_j^i &= \sum_{k=0}^{N-1}\sum_{l=0}^{N-1} c_{l+1}^{k+1}\omega^{(j-1)l}\omega^{(i-1)k} \\ &= 2 + \sum_{l=1}^{N-1}(-1)^l\omega^{(j-1)l} + \sum_{k=1}^{N-1}(-1)^k\omega^{(i-1)k} \\ &= \sum_{l=0}^{N-1}(-1)^l\omega^{(j-1)l} + \sum_{k=0}^{N-1}(-1)^k\omega^{(i-1)k}. \end{aligned} \tag{4.51}$$

These two properties of ω, the basic Nth root of unity, are needed for the following manipulation

$$\omega^{Nl} = e^{2\pi i N l/N} = (e^{2\pi i})^l = 1 \tag{4.52}$$

and

$$\omega^{ml} = e^{2\pi i m l/N} = (e^{\pi i})^l = (-1)^l, \tag{4.53}$$

where $N = 2m$. Now the sums in Eq. (4.51) can be written as

$$\sum_{l=0}^{N-1}(-1)^l\omega^{(j-1)l} = \sum_{l=0}^{N-1}\omega^{ml}\omega^{(j-1)l} = \sum_{l=0}^{N-1}\omega^{(m+j-1)l}, \tag{4.54}$$

which is a geometric sum and, assuming that $\omega^{m+j-1} \neq 1$, equals

$$\sum_{l=0}^{N-1}\omega^{(m+j-1)l} = \frac{1-\omega^{(m+j-1)N}}{1-\omega^{m+j-1}} = \frac{1-1}{1-\omega^{m+j-1}} = 0.$$

If $\omega^{m+j-1} = 1$, however, $m + j - 1 = N \Leftrightarrow j = m + 1$ and the sum becomes

$$\sum_{l=0}^{N-1}\omega^{(m+j-1)l} = \sum_{l=0}^{N-1} 1 = N.$$

Plugging in these two possible outcomes for the sums in Eq. (4.51) gives the possible eigenvalues

$$\lambda_j^i = \begin{cases} 2N & , \quad i = m+1,\ j = m+1 \\ N & , \quad i = m+1,\ j \neq m+1 \\ N & , \quad i \neq m+1,\ j = m+1 \\ 0 & , \quad i \neq m+1,\ j \neq m+1 \end{cases}. \tag{4.55}$$

If $i = m + 1$, it is possible to choose $N - 1$ values of j so that $j \neq m + 1$ and vice versa. Therefore, the multiplicity of the eigenvalue N is $2N - 2$, and the total number of non-zero eigenvalues of $\mathcal{B}$ is $2N - 1$.

Since the Δ matrix is the block matrix $\mathcal{B}$ multiplied by the scalar factor $\frac{h\omega^2}{2\pi^2}M$, the eigenvalues of Δ are those of $\mathcal{B}$ multiplied by the scalar factor and they share the same eigenvectors. M is the integration limit in Eq. (4.3) and thus $M \to \infty$. Therefore, all the non-zero eigenvalues of $\mathcal{B}$ correspond to infinitely large eigenvalues of Δ. As in 1-D, the Hamiltonian $\mathrm{H} = \mathrm{H}_0 + \Delta$ can be transformed to the basis of the eigenvectors of Δ, which gives

$$\mathrm{H} = \mathcal{V}_\mathrm{I}\mathcal{V}_1(\mathrm{H}_0' + \frac{h\omega^2}{2\pi^2}M\Lambda)(\mathcal{V}_\mathrm{I}\mathcal{V}_1)^{-1}, \tag{4.56}$$

where Λ is the diagonal matrix with the eigenvalues of $\mathcal{B}$ as its elements. As was reasoned in the 1-D case, adding the $2N - 1$ infinite diagonal elements forces the corresponding basis vectors to become eigenvectors of the Hamiltonian H and the eigenvalues of these eigenvectors are infinite. The rest of the eigenvectors, which are the reasonable eigenstates with finite eigenvalues, must obviously be orthogonal to

the eigenvectors corresponding to the infinite eigenvalues. Consequently, the addition of the Δ matrix to H_0, which is the Hamiltonian where the Gauss quadrature approximation is used, effectively removes the components to the directions of the Δ matrix eigenvectors with infinite eigenvalues from the eigenvectors of H_0. Therefore, those eigenvectors of H_0 that are almost orthogonal to the eigenvectors of Δ do not change significantly when the Δ matrix is added. It will be shown next that the eigenvectors of Δ that have infinite eigenvalues are analogous to the $\mathbf{x}$ eigenvector of the one dimensional Δ matrix, and thus a number of eigenvectors corresponding to the lowest eigenvalues of H_0 are approximately orthogonal to those eigenvectors of Δ.

4.3.3 Eigenvectors corresponding to non-zero eigenvalues

The eigenvectors of Δ are given by the columns of the matrix $\mathcal{W} = \mathcal{V}_\mathrm{I}\mathcal{V}_1$, where

$$\mathcal{V}_\mathrm{I} = \mathcal{V}_N(\mathrm{I}, \omega\mathrm{I}, \ldots, \omega^{N-1}\mathrm{I}) = \begin{bmatrix} \mathrm{I} & \mathrm{I} & \cdots & \mathrm{I} \\ \mathrm{I} & \omega\mathrm{I} & \cdots & \omega^{N-1}\mathrm{I} \\ \mathrm{I} & \omega^2\mathrm{I} & \cdots & \omega^{2(N-1)}\mathrm{I} \\ \vdots & \vdots & \cdots & \vdots \\ \mathrm{I} & \omega^{N-1}\mathrm{I} & \cdots & \omega^{(N-1)(N-1)}\mathrm{I} \end{bmatrix}$$

and

$$\mathcal{V}_1 = \mathrm{diag}(\mathrm{V}_1, \mathrm{V}_1, \ldots, \mathrm{V}_1),$$

where

$$\mathrm{V}_1 = \mathrm{V}(1, \omega, \ldots, \omega^{N-1}) = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & \omega & \cdots & \omega^{N-1} \\ 1 & \omega^2 & \cdots & \omega^{2(N-1)} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & \omega^{N-1} & \cdots & \omega^{(N-1)(N-1)} \end{bmatrix}.$$

The eigenvectors are not normalized, but that is not a problem because it is enough to show that the eigenvectors with non-zero eigenvalues are similar to the $\mathbf{x}$ eigenvector in the 1-D case. Performing the matrix multiplication block by block results in

$$\mathcal{W} = \mathcal{V}_\mathrm{I}\mathcal{V}_1 = \begin{bmatrix} \mathrm{V}_1 & \mathrm{V}_1 & \mathrm{V}_1 & \cdots & \mathrm{V}_1 \\ \mathrm{V}_1 & \omega\mathrm{V}_1 & \omega^2\mathrm{V}_1 & \cdots & \omega^{N-1}\mathrm{V}_1 \\ \mathrm{V}_1 & \omega^2\mathrm{V}_1 & \omega^4\mathrm{V}_1 & \cdots & \omega^{2(N-1)}\mathrm{V}_1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathrm{V}_1 & \omega^{N-1}\mathrm{V}_1 & \omega^{2(N-1)}\mathrm{V}_1 & \cdots & \omega^{(N-1)(N-1)}\mathrm{V}_1 \end{bmatrix}. \tag{4.57}$$

$\mathcal{W}$ is an $NN \times NN$ block matrix and can be rewritten as

$$\mathcal{W} = [\mathbf{w}_1 \quad \mathbf{w}_2 \quad \cdots \quad \mathbf{w}_N],$$

where

$$\mathbf{w}_i = \begin{bmatrix} \mathrm{V}_1 \\ \omega^{i-1}\mathrm{V}_1 \\ \omega^{2(i-1)}\mathrm{V}_1 \\ \vdots \\ \omega^{(N-1)(i-1)}\mathrm{V}_1 \end{bmatrix}$$

are the $NN \times N$ block vectors of $\mathcal{W}$. As can be verified from the diagonalized form of $\mathcal{B}$ in Eq. (4.49), the index i corresponds to the one in the eigenvalue formula of Eq. (4.51). Therefore, i in the eigenvalue formula is constant for each block vector $\mathbf{w}_i$

The block vector $\mathbf{w}_i$ can in turn be written using its component vectors as

$$\mathbf{w}_i = [\mathbf{v}_1^i \quad \mathbf{v}_2^i \quad \cdots \quad \mathbf{v}_N^i],$$

where the component vectors are

$$\mathbf{v}_j^i = \begin{bmatrix} \mathbf{x}_j \\ \omega^{i-1}\mathbf{x}_j \\ \omega^{2(i-1)}\mathbf{x}_j \\ \vdots \\ \omega^{(N-1)(i-1)}\mathbf{x}_j \end{bmatrix}, \quad \text{where } \mathbf{x}_j = \begin{bmatrix} 1 \\ \omega^{j-1} \\ \omega^{2(j-1)} \\ \vdots \\ \omega^{(N-1)(j-1)} \end{bmatrix}. \tag{4.58}$$

$\mathbf{v}_j^i$ are the eigenvectors of $\mathcal{B}$ with the corresponding eigenvalues λ_j^i given by Eq. (4.55). The eigenvector $\mathbf{v}_j^i$ can be written using its components as

$$\mathbf{v}_j^i = [v_1 \quad v_2 \quad \cdots \quad v_{N^2}].$$

The components of the eigenvector $\mathbf{v}_j^i$ correspond to the coefficients of the wave function expansion in Eq. (3.10) so that

$$v_k = c_{\mathbf{a}_k},$$

where

$$\mathbf{a}_k = (a_{x,k},\, a_{y,k}), \quad a_{x,k}, a_{y,k} \in \left\{ -\frac{N-1}{2},\, \frac{N-1}{2} + 1,\, \ldots,\, \frac{N-1}{2} \right\}.$$

The two dimensional Hamiltonian matrix is formed so that a_x runs through the possible values and a_y stays constant. Then a_y is increased by one and a_x runs again, and so on. The index k can thus be defined as

$$k = \left(a_y + \frac{N-1}{2} \right) N + a_x + \frac{N-1}{2} + 1.$$

Consequently, the y-coordinate is constant for the N successive components of $\mathbf{v}_j^i$ for which

$$k \in [iN + 1, iN + N], \quad i \in \{0, 1, \ldots, N-1\}.$$

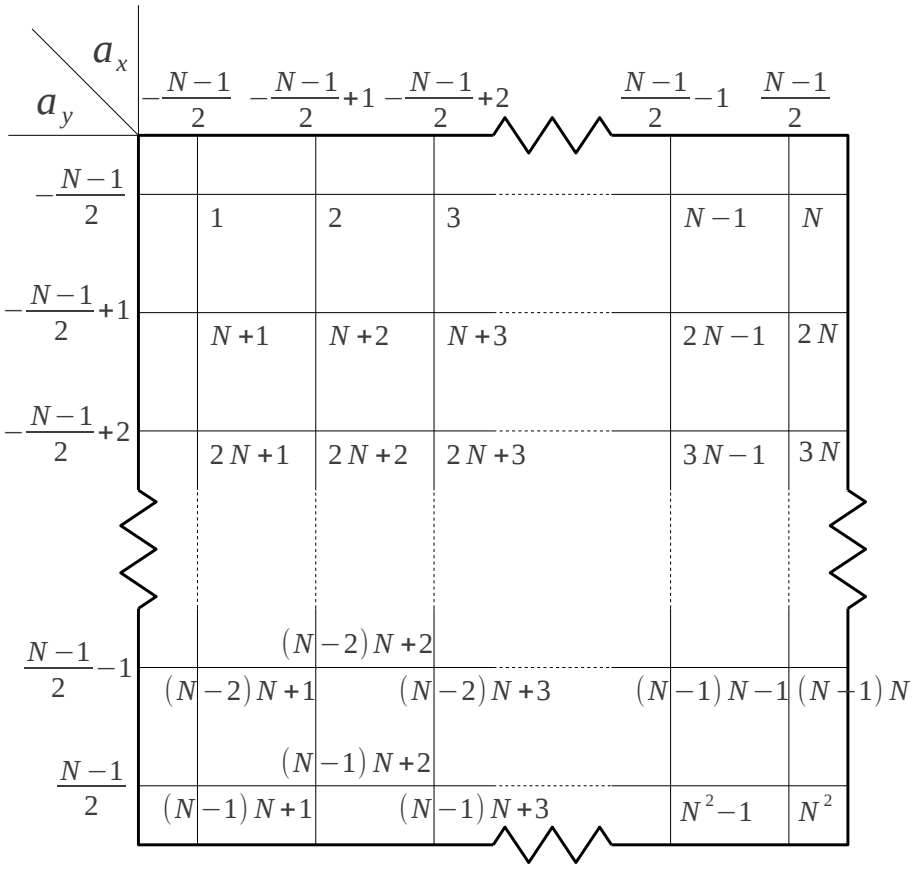

Figure 4.3: The indexing of the mesh points in the Lagrange mesh method. The numbers inside the mesh denote the k indices of the eigenvector components.

The schematic drawing of the mesh in Fig. 4.3 illustrates the indexing more clearly.

If i is chosen so that $i = m + 1$, the component vectors become

$$\mathbf{v}_j^{m+1} = \begin{bmatrix} \mathbf{x}_j \\ \omega^m \mathbf{x}_j \\ \omega^{2m} \mathbf{x}_j \\ \vdots \\ \omega^{(N-1)m} \mathbf{x}_j \end{bmatrix} = \begin{bmatrix} \mathbf{x}_j \\ -\mathbf{x}_j \\ \mathbf{x}_j \\ \vdots \\ -\mathbf{x}_j \end{bmatrix} . \tag{4.59}$$

$\mathbf{v}_j^{m+1}$ is an eigenvector of $\mathcal{B}$ with an eigenvalue $\lambda_j^{m+1} = N$, assuming that $j \neq m+1$. The mesh of basis function coefficients corresponding to the eigenvector $\mathbf{v}_j^{m+1}$ is shown in Fig. 4.4. It has been obtained by placing the elements of the vector $\mathbf{v}_j^{m+1}$ at the mesh points where the corresponding Lagrange functions are centered. For each x-coordinate, the magnitude of the coefficient is constant but the sign alternates in the y-direction. The eigenvector $\mathbf{v}_j^{m+1}$ is clearly analogous to the $\mathbf{x}$ eigenvector of the Δ matrix in the 1-D case, which given by Eq. (4.13). The only difference is the coordinate of the extra dimension which can have N distinct values. If $\mathbf{u}$ is an

	x_1	x_2	x_3	x_{N-1}	x_N
	$-x_1$	$-x_2$	$-x_3$	$-x_{N-1}$	$-x_N$
	x_1	x_2	x_3	x_{N-1}	x_N
	x_1	x_2	x_3	x_{N-1}	x_N
	$-x_1$	$-x_2$	$-x_3$	$-x_{N-1}$	$-x_N$

Figure 4.4: The mesh of Lagrange basis function coefficients corresponding to the eigenvector $\mathbf{v}_j^{m+1}$. x_k are the components of the vector $\mathbf{x}_j$, i.e. $\mathbf{x}_j = [x_1 \quad x_2 \quad \cdots \quad x_N]^{\mathrm{T}}$ and $x_k = \omega^{(k-1)(j-1)}$.

eigenstate of the Hamiltonian H_0, the inner product of $\mathbf{u}$ with $\mathbf{v}_j^i$ is

$$\begin{aligned} \mathbf{u}^{\mathrm{T}}\mathbf{v}_j^i &= \sum_{k=1}^{N^2} u_k v_k \\ &= \sum_{a_x=-\frac{N-1}{2}}^{\frac{N-1}{2}} \sum_{a_y=-\frac{N-1}{2}}^{\frac{N-1}{2}} d_{(a_x,a_y)} c_{(a_x,a_y)}, \end{aligned} \tag{4.60}$$

where $d_{(a_x,a_y)}$ are the basis function coefficients corresponding to the components of the eigenvector $\mathbf{u}$ and $c_{(a_x,a_y)}$ are those corresponding to $\mathbf{v}_j^i$. The inner product is thus separable to sums where either x or y-coordinate is constant. Now the inner product of $\mathbf{u}$ and $\mathbf{v}_j^{m+1}$ can be obtained by reading the coefficients $c_{(a_x,a_y)}$ from the coefficient mesh in Fig. 4.4, and it becomes

$$\begin{aligned} \mathbf{u}^{\mathrm{T}}\mathbf{v}_j^{m+1} &= \sum_{b_x=1}^{N} x_{b_x} \sum_{b_y=1}^{N} (-1)^{b_y-1} d_{\left(t(b_x),t(b_y)\right)} \\ &= \sum_{b_x=1}^{N} x_{b_x} \sum_{b_y=1}^{N/2} \left[d_{\left(t(b_x),t(2b_y-1)\right)} - d_{\left(t(b_x),t(2b_y)\right)} \right] \end{aligned}$$

where $t(x) = x - \frac{N-1}{2} - 1$ is a coordinate translation to simplify the notation. The sum over b_y is essentially the same as the sum in Eq. (A.4), which gives the inner product in the 1-D case. Therefore, the sum could be manipulated in the same way as in Appendix A to obtain a sum of wave function derivatives in the y-direction,

and relate that to an integral of the derivative over the length of the simulation area. The magnitude of the inner product $\mathbf{u}^{\mathrm{T}}\mathbf{v}_j^{m+1}$ is thus bounded so that $|\mathbf{u}^{\mathrm{T}}\mathbf{v}_j^{m+1}| \to 0$ when the mesh spacing $h \to 0$ and should be small for the lowest eigenstates of H_0 for a reasonable value of h.

	y_1	$-y_1$	y_1	y_1	$-y_1$
	y_2	$-y_2$	y_2	y_2	$-y_2$
	y_3	$-y_3$	y_3	y_3	$-y_3$
	y_{N-1}	$-y_{N-1}$	y_{N-1}	y_{N-1}	$-y_{N-1}$
	y_N	$-y_N$	y_N	y_N	$-y_N$

Figure 4.5: The mesh of Lagrange basis function coefficients corresponding to the eigenvector $\mathbf{v}_{m+1}^i$. $y_k = \omega^{(k-1)(i-1)}$.

If $j = m+1$ instead of $i = m+1$, the eigenvectors are

$$\mathbf{v}_{m+1}^i = \begin{bmatrix} \mathbf{x}_{m+1} \\ \omega^{i-1}\mathbf{x}_{m+1} \\ \omega^{2(i-1)}\mathbf{x}_{m+1} \\ \vdots \\ \omega^{(N-1)(i-1)}\mathbf{x}_{m+1} \end{bmatrix} = \begin{bmatrix} y_1\,\mathbf{x}_{m+1} \\ y_2\,\mathbf{x}_{m+1} \\ y_3\,\mathbf{x}_{m+1} \\ \vdots \\ y_N\,\mathbf{x}_{m+1} \end{bmatrix}, \tag{4.61}$$

where

$$\mathbf{x}_{m+1} = \begin{bmatrix} 1 \\ \omega^m \\ \omega^{2m} \\ \vdots \\ \omega^{(N-1)m} \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 1 \\ \vdots \\ -1 \end{bmatrix}. \tag{4.62}$$

$\mathbf{v}_{m+1}^i$ is an eigenvector of $\mathcal{B}$ with an eigenvalue $\lambda_{m+1}^i = N$, assuming that $i \neq m+1$. The mesh of the basis function coefficients corresponding to $\mathbf{v}_{m+1}^i$ is shown in Fig. 4.5. Now the sign alternates in the x-direction and the inner product with a eigenvector of the Hamiltonian H_0 would lead to an integral of the x-derivative of the wave function.

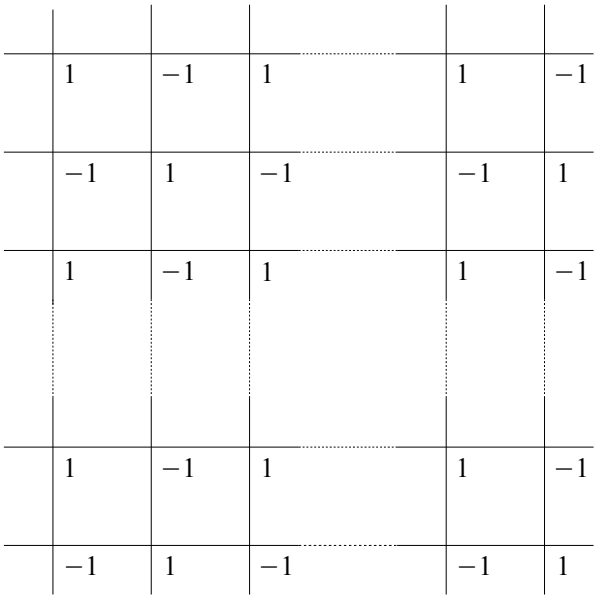

Figure 4.6: The mesh of Lagrange basis function coefficients corresponding to the eigenvector $\mathbf{v}_{m+1}^{m+1}$.

There is still the case $i = m + 1$ and $j = m + 1$ left. The eigenvector for these index values is

$$\mathbf{v}_{m+1}^{m+1} = \begin{bmatrix} \mathbf{x}_{m+1} \\ \omega^{m}\mathbf{x}_{m+1} \\ \omega^{2m}\mathbf{x}_{m+1} \\ \vdots \\ \omega^{(N-1)m}\mathbf{x}_{m+1} \end{bmatrix} = \begin{bmatrix} \mathbf{x}_{m+1} \\ -\mathbf{x}_{m+1} \\ \mathbf{x}_{m+1} \\ \vdots \\ -\mathbf{x}_{m+1} \end{bmatrix}, \tag{4.63}$$

where $\mathbf{x}_{m+1}$ is given by Eq. (4.62). $\mathbf{v}_{m+1}^{m+1}$ is an eigenvector of $\mathcal{B}$ with an eigenvalue $\lambda_{m+1}^{m+1} = 2N$. The coefficient mesh is shown in Fig. 4.6. Now the sign alternates in both x and y-direction and the inner product with a eigenvector of the Hamiltonian H_0 would lead to an integral of the combined x and y-derivative of the wave function over the whole simulation area.

The previous analysis of the eigenvectors of the Δ matrix shows that the eigenvectors corresponding to the infinite eigenvalues are analogous to the $\mathbf{x}$ eigenvector of the Δ matrix in 1-D case. Therefore, the same arguments can be used to show that they are approximately orthogonal to a number of lowest eigenvectors of the Hamiltonian H_0. Consequently, the lowest eigenvectors of H_0 stay approximately the same when the components of $\mathbf{v}_j^{m+1}$, $\mathbf{v}_{m+1}^{i}$ and $\mathbf{v}_{m+1}^{m+1}$ are removed from them. All this reasoning leads to the fact that the addition of the Δ matrix to H_0 does not notably affect the lowest eigenstates of H_0 if the mesh spacing h is small enough and the size of the simulation area is sufficient. The suitable values can be found by performing numerical convergence tests.

4.3.4 Eliminating the effect of the approximation

The analysis that is performed in this chapter shows that the error caused by the Gauss quadrature approximation decreases as the mesh spacing h is decreased and the size of the simulation area is increased. According to the convergence tests performed in Section 5.1, the energies of the single-particle states in a parabolic potential converge very fast towards the exact energies of the Fock-Darwin states. Therefore, the error caused by the Gauss quadrature approximation seems to be extremely small in the range of reasonable mesh spacings. However, since the Gauss quadrature approximation breaks the variational principle, as was noted in Section 4.2.3, the energies given by the Lagrange method can be smaller than correct ones. Consequently, one cannot, for example, find the optimal size of the simulation area when the number of mesh points is fixed by minimizing the energies of the single-particle states. If it is important that the variational principle is preserved, the following method can be used to eliminate the effect of the Gauss quadrature approximation.

The analysis of the effect caused by the Gauss quadrature approximation leads directly into a way for completely eliminating that effect. As a short revision, the difference between the Hamiltonians that are calculated with and without the Gauss quadrature is given by the Δ matrix, which in 2-D is as in Eq. (4.28). The Δ matrix was shown to have $2N-1$ infinite eigenvalues, and the rest of them are zero. If the whole Hamiltonian, $\mathrm{H} = \mathrm{H}_0 + \Delta$, is transformed to the basis of the eigenvectors of Δ, the difference between H and H_0 resides in $2N-1$ diagonal elements, and the differences are the infinite eigenvalues of Δ. Adding these infinite diagonal elements to H_0 forces the corresponding basis vectors, i.e. the corresponding eigenvectors of Δ, to become eigenvectors of H with an infinite eigenvalue. Since the rest of the eigenvectors must be orthogonal to these, the eigenvectors of H_0 are altered so that any component in the direction of the eigenvectors with the infinite eigenvalues are removed from them.

The addition of the Δ matrix to the Hamiltonian H_0, where the potential matrix elements are approximated using the Gauss quadrature, seems to have the same effect as if the components of the eigenvectors of Δ that have the infinite eigenvalues would be removed from the Hamiltonian. This offers a numerically stable way to compute the eigenvalues of H. The Lagrange mesh Hamiltonian H_0 is first transformed to the basis of the eigenvectors of the Δ matrix. In 2-D, the eigenvectors of Δ are the columns of the block matrix $\mathcal{W}$, which is given by Eq. (4.57). It is assumed that N, the number of mesh points in one direction, is even. Since $\mathcal{W}$ consists of the orthogonal eigenvectors of the symmetric matrix Δ, normalization of the eigenvectors simplifies the inversion of $\mathcal{W}$, because $\hat{\mathcal{W}}^{-1} = \hat{\mathcal{W}}^*$ if the eigenvectors are normalized. It is easy to see that the normalization factor of each vector is simply $1/N$, and thus $\hat{\mathcal{W}} = \mathcal{W}/N$. The Hamiltonian can be transformed by performing two matrix products

$$\mathrm{H}_0' = \mathcal{W}^* \, \mathrm{H}_0 \, \mathcal{W}, \tag{4.64}$$

where $\mathcal{W}$ is normalized. Now, it is possible to represent the original Hamiltonian

with respect to the orthonormal basis given by $\mathcal{W}$

$$\mathrm{H}_0 = \mathcal{W}\,\mathrm{H}_0'\,\mathcal{W}^*.$$

As the addition of the Δ matrix to H_0 effectively removes the components to the directions of the eigenvectors of Δ that have infinite eigenvalues, the same effect should be achieved by manually removing the corresponding basis vectors from $\mathcal{W}$. The infinite eigenvalues of Δ are the non-zero eigenvalues of $\mathcal{B}$, which are given in Eq. (4.55). The eigenvalue λ_j^i of $\mathcal{B}$ is zero if $i \neq N/2+1$ and $j \neq N/2+1$. If k denotes the index of a column vector in Δ, it can be related to the indices i and j by

$$k = (i-1)N + j.$$

The new Hamiltonian is built by taking into account only the eigenvectors in the basis $\mathcal{W}$ that have a zero eigenvalue, i.e. those columns of $\mathcal{W}$ for which the index k is such that $i \neq N/2+1$ and $j \neq N/2+1$. Consequently, the fixed Hamiltonian can be written as

$$\mathrm{H}_{fix} = \tilde{\mathcal{W}}\,\tilde{\mathrm{H}}_0'\,\tilde{\mathcal{W}}^*, \tag{4.65}$$

where $\tilde{\mathcal{W}}$ consists of those columns of $\mathcal{W}$ for which $i \neq N/2+1$ and $j \neq N/2+1$, and $\tilde{\mathrm{H}}_0'$ consists of columns and rows of H_0' with the same condition for i and j. The eigenvalues and eigenvectors of the Hamiltonian H_{fix} should be the same as those of $\mathrm{H} = \mathrm{H}_0 + \Delta$, except that the infinite eigenvalues are replaced by zeros, because those components have been removed. This method to eliminate the effect of the Gauss quadrature approximation is tested in Section 5.1, and it seems to work perfectly.

4.4 Beyond the simple parabolic potential

The analysis of the Gauss quadrature approximation in the case of a single parabolic potential led to an excellent understanding of how the approximation affects the eigenstates and their energies. Even a method for eliminating the effect completely was derived. However, there is no point in solving the single-particle states in a single parabolic potential using the Lagrange mesh method because the states are analytically solvable. These analytically solved states are the so called Fock-Darwin states. The interesting applications of the Lagrange mesh method are in more complex systems such as a double quantum dot consisting of two parabolic minima, which is studied in this thesis, or even a system of two double quantum dots consisting of four parabolic minima. It is thus obvious that study of the Gauss quadrature approximation must be extended to more complicated potentials.

As the exact form of the electrostatic potential in semiconductor quantum dots is unknown and varies somewhat from device to device, the potential is usually approximated to be parabolic at the minimum of the potential. The most straightforward way to extend this approximation to systems consisting of two or more quantum dots is to make the potential piecewise parabolic. In that case, the potential can be defined to be the minimum of the parabolic functions centred at different points. This piecewise definition produces sharp edges at the intersections of the parabolic

minima, which means that the derivatives of the potential are not continuous at the intersections. The derivatives, or at least the first one, can be made continuous by rounding the intersections somehow. An extremely simple way for doing this is presented in Section 5.1. One could also consider using a fourth degree polynomial, which can be tuned to have two identical minima, to model a double quantum dot. Building a sensible model for more than two quantum dots from a higher degree polynomial becomes a tricky task, however.

In this thesis, the double quantum dot potentials are modelled as piecewise parabolic potentials. The convergence rates of single-particle energies and the two-particle ground state energy for rounded and non-rounded double quantum dot potentials are compared in Section 5.1. The convergence of single-particle states in double quantum dot modelled by a fourth degree polynomial is also checked for comparison. These tests show that the convergence rates are substantially faster for the rounded piecewise parabolic potential than for the non-rounded one. Furthermore, the convergence of the single-particle energies in a fourth degree polynomial potential are in their own league when compared to the piecewise parabolic potentials. It would seem that the error due to the Gauss quadrature approximation is extremely small for the fourth degree polynomial, and this raises the question, whether the results that were presented for the parabolic potential are generalizable for higher degree polynomials.

4.4.1 Matrix elements for arbitrary degree polynomials

Suppose that the two dimensional potential is a polynomial of arbitrary degree

$$V(x,y) = \sum_{i=1}^{m} a_i x^{n_x(i)} y^{n_y(i)}, \tag{4.66}$$

where $n_x(i) \in \mathbb{N}$ and $n_y(i) \in \mathbb{N}$. The matrix elements for such a potential are given by

$$\begin{aligned} \langle L_{\mathbf{a}'}|V(x,y)|L_{\mathbf{a}}\rangle &= \sum_{i=1}^{m} a_i \langle L_{\mathbf{a}'}|x^{n_x(i)} y^{n_y(i)}|L_{\mathbf{a}}\rangle \\ &= \sum_{i=1}^{m} a_i \langle L_{a'_x}|x^{n_x(i)}|L_{a_x}\rangle \langle L_{a'_y}|y^{n_y(i)}|L_{a_y}\rangle \,. \end{aligned} \tag{4.67}$$

As each term separates with respect to the position variables x and y, and the integrals over the variables are identical, the matrix elements are easily obtained if one finds a simple expression for the integrals of type

$$R_{a'a}(n) = \langle L_{a'}|x^n|L_a\rangle = \frac{1}{h}\int_{-\infty}^{\infty} x^n \frac{\sin[\frac{\pi}{h}(x-x_{a'})]\sin[\frac{\pi}{h}(x-x_a)]}{\frac{\pi}{h}(x-x_{a'})\frac{\pi}{h}(x-x_a)}\,. \tag{4.68}$$

$R_{a'a}(2)$ corresponds to the integral that was needed in the calculation of the matrix elements for the parabolic potential. Now the integral will be solved for arbitrary n using recursion.

First, the integral is simplified by making a change of variable $x = ht$,

$$R_{ab}(n) = \frac{h^n}{\pi^2} \int_{-M'}^{M'} t^n \, \frac{\sin[\pi(t-a)] \sin[\pi(t-b)]}{(t-a)(t-b)} dt = \frac{h^n}{\pi^2} F(n,a,b)\,. \tag{4.69}$$

Here, $a = x'_a/h$ and $b = x_a/h$, and the integration limit $M' = M/h$. The limit $M \to \infty$ will be taken in the end. A recursive form for the integral $F(n,a,b)$ can be found by making the following manipulations

$$\begin{aligned}
F(n,a,b) &= \int_{-M'}^{M'} t^{n-2} \, (t-a+a) \frac{1}{t-a} (t-b+b) \frac{1}{t-b} \\
&\quad \times \sin[\pi(t-a)] \sin[\pi(t-b)] \, dt \\
&= \int_{-M'}^{M'} t^{n-2} \left(1 + \frac{a}{t-a} + \frac{b}{t-b} + \frac{ab}{(t-a)(t-b)} \right) \\
&\quad \times \sin[\pi(t-a)] \sin[\pi(t-b)] \, dt \\
&= \int_{-M'}^{M'} t^{n-2} \, \sin[\pi(t-a)] \sin[\pi(t-b)] \, dt \\
&\quad + a \int_{-M'}^{M'} t^{n-2} \, \frac{\sin[\pi(t-a)] \sin[\pi(t-b)]}{t-a} \, dt \\
&\quad + b \int_{-M'}^{M'} t^{n-2} \, \frac{\sin[\pi(t-a)] \sin[\pi(t-b)]}{t-b} \, dt \\
&\quad + ab \int_{-M'}^{M'} t^{n-2} \, \frac{\sin[\pi(t-a)] \sin[\pi(t-b)]}{(t-a)(t-b)} \, dt \\
&= ab\, F(n-2,a,b) + a\, G(n-2,a) + b\, G(n-2,b) + L(n-2).
\end{aligned} \tag{4.70}$$

The recursion formula is clearly valid only for $n \geq 2$. The base cases $n = 0$ and $n = 1$ can be easily integrated to obtain $F(0,a,b) = \pi^2 \delta_{ab}$ and $F(1,a,b) = \pi^2 a \delta_{ab}$. A recursion formula can be derived also for $G(n-2,a)$ and $G(n-2,b)$, and it becomes

$$\begin{aligned}
G(n,a) &= \int_{-M'}^{M'} t^n \, \frac{\sin[\pi(t-a)] \sin[\pi(t-b)]}{t-a} \, dt \\
&= \int_{-M'}^{M'} t^{n-1} \, \sin[\pi(t-a)] \sin[\pi(t-b)] \, dt \\
&\quad + a \int_{-M'}^{M'} t^{n-1} \, \frac{\sin[\pi(t-a)] \sin[\pi(t-b)]}{t-a} \, dt \\
&= a\, G(n-1,a) + L(n-1).
\end{aligned} \tag{4.71}$$

The base case $G(0,a)$ has been solved in Section 4.1, and it gives $G(0,a) = 0$. Due to its simple form, the recursion relation of $G(n,a)$ can be written as a sum

$$G(n,a) = \sum_{i=0}^{n-1} a^{n-1-i} L(i). \tag{4.72}$$

This expression for $G(n,a)$ can be used to simplify the the recursion relation of $F(n,a,b)$ in Eq. (4.70), and it results in

$$F(n,a,b) = ab\,F(n-2,a,b) + \sum_{i=0}^{n-3}(a^{n-2-i} + b^{n-2-i})L(i) + L(n-2). \tag{4.73}$$

The integrals $L(n)$ in the recursion relation are still unknown and will be solved next. The product of two sine functions can be expressed as a sum of cosine functions using a trigonometric identity

$$\begin{aligned} L(n) =& \int_{-M'}^{M'} t^n \sin[\pi(t-a)]\sin[\pi(t-b)]dt \\ =& \int_{-M'}^{M'} t^n \frac{\cos[\pi(b-a)] - \cos[2\pi t - \pi(a+b)]}{2} dt \\ =& \frac{(-1)^{a-b}}{2}\int_{-M'}^{M'} t^n dt - \frac{(-1)^{a+b}}{2}\int_{-M'}^{M'} t^n\cos(2\pi t)dt. \end{aligned} \tag{4.74}$$

Since a and b are unscaled mesh points, and thus defined as a_x in Eq. (3.7), it is obvious that $b-a \in \mathbb{Z}$ and $a+b \in \mathbb{Z}$. Consequently, the relation above has been simplified using $\cos[\pi(b-a)] = (-1)^{b-a}$ and $\cos[2\pi t - \pi(a+b)] = (-1)^{a+b}\cos(2\pi t)$. The integrands in both integrals of Eq. (4.74) are odd when n is odd, and therefore, $L(n) = 0$ when n is odd. The integral $L(n)$ has thus two possible outcomes

$$L(n) = \begin{cases} 0 & \text{, n odd} \\ (-1)^{a-b}\left[\frac{1}{n+1}M'^{n+1} - \frac{H(n)}{2}\right] & \text{, n even.} \end{cases} \tag{4.75}$$

Since the integration limit M' is taken into infinity at the end, the integral $L(n)$ diverges and only the asymptotic behaviour of $L(n)$ as a function of M' is interesting. One can partial integrate $H(n)$ to cast it into recurrence form

$$\begin{aligned} H(n) =& \int_{-M'}^{M'} t^n\cos(2\pi t)dt \\ =& \frac{1}{\pi}M'^n\sin(2\pi M') + \frac{n}{2\pi^2}M'^{n-1}\cos(2\pi M') - \frac{n(n-1)}{4\pi^2}H(n-2). \end{aligned} \tag{4.76}$$

From this recurrence form, it is easy to see that $H(n) = O(M'^n)$, where $O(M'^n)$ is the big O notation and means that

$$H(n) = O(M'^n) \Rightarrow |H(n)| \leq C|M'^n| \quad \text{for all } M' > M'_0,$$

where C is some positive constant. Therefore, when M' is larger than the threshold value M'_0, the growth of $H(n)$ as a function of M' can be described by the leading term M'^n. $L(n)$ can now be written as

$$L(n) = \begin{cases} 0 & \text{, n odd} \\ (-1)^{a-b}\left[\frac{1}{n+1}M'^{n+1} + O(M'^n)\right] & \text{, n even.} \end{cases} \tag{4.77}$$

Using the derived expression for $L(n)$, the recursion relation of $F(n,a,b)$ in Eq. (4.73) can be further simplified to obtain

$$F(n,a,b) = ab\,F(n-2,a,b) + (-1)^{a-b}\begin{cases} \displaystyle\sum_{i=0}^{(n-4)/2} (a^{n-2-2i}+b^{n-2-2i})\,\frac{M'^{2i+1}}{2i+1} + \frac{M'^{n-1}}{n-1} & \text{, n even} \\ \displaystyle\sum_{i=0}^{(n-3)/2} (a^{n-2-2i}+b^{n-2-2i})\,\frac{M'^{2i+1}}{2i+1} & \text{, n odd}\,. \end{cases} \tag{4.78}$$

The notation is somewhat sloppy here because the limit $M' \to \infty$ has already been anticipated and thus only the leading term in $L(n)$ has been considered. One could now argue that only the term with the highest degree of M' is relevant in $F(n,a,b)$ when the limit $M' \to \infty$ is taken, but since $F(n,a,b)$ describes the potential matrix elements, one must be careful not to remove any dependency on a and b. a and b define the position of the element in the matrix, and thus they are extremely important when the eigenvalues and eigenvectors of the matrix are calculated.

4.4.2 Example cases and general hypothesis

The recurrence relation for $F(n,a,b)$ in Eq. (4.78) together with Eq. (4.69) can be used to calculate exact potential matrix elements for a polynomial potential of arbitrary degree. However, the same problem that was noticed in the case of a parabolic potential, namely the divergence of the elements, persist for any polynomial that has a degree higher than one. Therefore, the exact potential matrix elements cannot be used in the numerical calculations and one has to rely on the elements that are obtained using the Gauss quadrature approximation. The Gauss quadrature was found to be a good approximation in the case of the parabolic potential, because the error it causes decreases fast as the mesh spacing h is decreased and the size of the simulation area is increased. To understand how the Gauss quadrature approximation affects electron states in a polynomial potential of degree $n > 2$, two examples of polynomial potentials in 1-D, $n = 4$ and $n = 6$, are analysed in this section.

The potential matrix elements of the fourth and sixth degree polynomial potentials, $V_4(x) = C_4x^4$ and $V_6(x) = C_6x^6$, are

$$V_{4,ab} = C_4\langle L_a|x^4|L_b\rangle = C_4R_{ab}(4) = \frac{C_4h^4}{\pi^2}F(4,a,b)$$
$$V_{6,ab} = \frac{C_6h^6}{\pi^2}F(6,a,b).$$

Using the recursion relation of Eq. (4.78), the elements become

$$\begin{aligned} V_{4,ab} &= \frac{C_4 h^4}{\pi^2}\left[\pi^2 a^2 b^2 \delta_{ab} + (-1)^{a-b}\left(\frac{M'^3}{3} + (a^2+ab+b^2)M'\right)\right] \\ &= V_4(x_a)\delta_{ab} + \Delta_{4,ab} \\ V_{6,ab} &= \frac{C_6 h^6}{\pi^2}\left[\pi^2 a^3 b^3 \delta_{ab} + (-1)^{a-b}\left(\frac{M'^5}{5} + (a^2+ab+b^2)\frac{M'^3}{3} + \right.\right. \\ &\qquad \left.\left. + (a^4+a^3b+a^2b^2+ab^3+b^4)M'\right)\right] \\ &= V_6(x_a)\delta_{ab} + \Delta_{6,ab}\,, \end{aligned}$$

where the delta matrices

$$\Delta_{4,ab} = (-1)^{a-b}\,\frac{C_4 h^4}{\pi^2}\left(\frac{M'^3}{3} + (a^2+ab+b^2)\,M'\right) \tag{4.79}$$

and

$$\begin{aligned} \Delta_{6,ab} = (-1)^{a-b}\,\frac{C_6 h^6}{\pi^2}\Bigg(&\frac{M'^5}{5} + (a^2+ab+b^2)\,\frac{M'^3}{3} + \\ &+ (a^4+a^3b+a^2b^2+ab^3+b^4)\,M'\Bigg) \end{aligned} \tag{4.80}$$

give the differences between the exact matrix elements and the matrix elements that are obtained using the Gauss quadrature. If the addition of these matrices leaves the lowest eigenvalues and eigenvectors of the Hamiltonian almost untouched, as was the case for the parabolic potential, the Gauss quadrature gives a good approximation for the potential matrix elements. To investigate the effect of the addition of Δ matrices, their eigenvalues and eigenvectors must be solved.

The difference matrices for the fourth and sixth degree polynomial potentials are much more complicated than that for the second degree polynomial. Therefore, the eigenvalues and eigenvectors of a general $N \times N$ Δ matrix cannot be solved in these cases, or at least it would require substantial amount of extra work. Instead, the eigenvalues and eigenvectors are calculated for 8×8 matrices, for which $a, b \in \{-7/2, -5/2, -3/2, \ldots, 7/2\}$. The eigensystems were solved analytically using Mathematica, carrying the M' in the calculations until the limit $M' \to \infty$ was taken in the end. Interestingly, the difference matrix for the fourth degree polynomial, Δ_4, has two infinite eigenvalues and the other 6 eigenvalues are zeros. Δ_4 has thus one extra infinite eigenvalue compared to Δ_2, which is the difference matrix for the second degree polynomial. The eigenvectors of the 8×8 Δ_4 matrix that correspond to the infinite eigenvalues are

$$\mathbf{x}_1 = \frac{1}{K_1}\,[1 \quad -1 \quad 1 \cdots 1 \quad -1]^{\mathrm{T}} \tag{4.81}$$

and

$$\mathbf{x}_2 = \frac{1}{K_2}\,[-7/2 \quad 5/2 \quad -3/2 \cdots 5/2 \quad -7/2]^{\mathrm{T}}\,. \tag{4.82}$$

K_1 and K_2 are the normalization coefficients, which are given by

$$K_1 = \sqrt{8}, \quad K_2 = \sqrt{\sum_{i=1}^{8} \left(-\frac{7}{2} + i - 1\right)^2}.$$

The first eigenvector, $\mathbf{x}_1$, is equal to the eigenvector that corresponds to the infinite eigenvalue of the Δ_2 matrix. Solving the eigensystem of Δ_6 gives three infinite eigenvalues, and $\mathbf{x}_1$ and $\mathbf{x}_2$ are two of the eigenvectors. The third eigenvector is

$$\mathbf{x}_3 = \frac{1}{K_3} \begin{bmatrix} 7 & -1 & -3 & 5 & -5 & 3 & 1 & -7 \end{bmatrix}^{\mathrm{T}}, \tag{4.83}$$

which can also be written as

$$\mathbf{x}_3 = \frac{1}{K_3} \left(\tilde{\mathbf{x}}_3 - (\mathbf{x}_1 \cdot \tilde{\mathbf{x}}_3) \mathbf{x}_1 \right), \tag{4.84}$$

where the vector $\tilde{\mathbf{x}}_3$ is

$$\tilde{\mathbf{x}}_3 = \frac{1}{\tilde{K}_3} \left[\left(\frac{7}{2}\right)^2 \quad -\left(\frac{5}{2}\right)^2 \quad \left(\frac{3}{2}\right)^2 \cdots \left(\frac{5}{2}\right)^2 \quad -\left(\frac{7}{2}\right)^2 \right]^{\mathrm{T}}. \tag{4.85}$$

The normalization coefficient $\tilde{K}_3$ is

$$\tilde{K}_3 = \sqrt{\sum_{i=1}^{8} \left(-\frac{7}{2} + i - 1\right)^4}.$$

The eigenvectors of Δ_4 and Δ_6 matrices of size 8×8 are so well structured that it is easy to anticipate the form of the eigenvectors of general $N \times N$ matrices. It seems that the general eigenvectors can be written as

$$\mathbf{x}_1 = \frac{1}{K_1} \begin{bmatrix} 1 & -1 & 1 \cdots 1 & (-1)^{N-1} \end{bmatrix}, \tag{4.86}$$

$$\mathbf{x}_2 = \frac{1}{K_2} \begin{bmatrix} a_1 & -a_2 & a_3 \cdots a_{N-1} & (-1)^{N-1} a_N \end{bmatrix}, \tag{4.87}$$

$$\mathbf{x}_3 = \frac{1}{K_3} \left(\tilde{\mathbf{x}}_3 - (\mathbf{x}_1 \cdot \tilde{\mathbf{x}}_3) \mathbf{x}_1 \right), \tag{4.88}$$

where

$$\tilde{\mathbf{x}}_3 = \frac{1}{\tilde{K}_3} \begin{bmatrix} a_1^2 & -a_2^2 & a_3^2 \cdots a_{N-1}^2 & (-1)^{N-1} a_N^2 \end{bmatrix}. \tag{4.89}$$

Here, a_i are the unscaled mesh points of the Lagrange mesh given by

$$a_i = -\frac{N+1}{2} + i\,. \tag{4.90}$$

The normalization coefficients are now

$$K_1 = \sqrt{N}, \quad K_2 = \sqrt{\sum_{i=1}^{N} a_i^2}, \quad \tilde{K}_3 = \sqrt{\sum_{i=1}^{N} a_i^4}\,. \tag{4.91}$$

The formula for the eigenvector $\mathbf{x}_3$ in Eq. (4.88) may need some explanation. The vector $\tilde{\mathbf{x}}_3$ was found to be a linear combination of the eigenvectors $\mathbf{x}_1$ and $\mathbf{x}_3$, and thus it can be written

$$\tilde{\mathbf{x}}_3 = \alpha\mathbf{x}_1 + \beta\mathbf{x}_3\,,$$

where $\alpha = \mathbf{x}_1 \cdot \tilde{\mathbf{x}}_\mathbf{3}$ and $\beta = \mathbf{x}_3 \cdot \tilde{\mathbf{x}}_\mathbf{3}$. Since $\tilde{\mathbf{x}}_3$ is normalized, the condition

$$\alpha^2 + \beta^2 = 1$$

is required. From these relations, one can derive

$$\mathbf{x}_3 = \frac{1}{\sqrt{1 - (\mathbf{x}_1 \cdot \tilde{\mathbf{x}}_\mathbf{3})^2}}\left(\tilde{\mathbf{x}}_3 - (\mathbf{x}_1 \cdot \tilde{\mathbf{x}}_\mathbf{3})\mathbf{x}_1\right), \tag{4.92}$$

which also shows that the normalization coefficient K_3 must be

$$K_3 = \sqrt{1 - (\mathbf{x}_1 \cdot \tilde{\mathbf{x}}_\mathbf{3})^2}\,. \tag{4.93}$$

It should be noted that this is not a mathematical proof that eigenvectors of the $N \times N$ matrices Δ_4 and Δ_6 are the ones shown here, as their form has merely been derived from the special case of 8×8 matrices. However, the derived formulas for the eigenvectors were found to agree with various choices of matrix sizes, including $N \in \{6, 7, 8, 9\}$. It is thus safe to assume that these formulas are correct.

As was reasoned in the case of the parabolic potential, adding the difference matrix Δ to the Hamiltonian matrix H_0 with the Gauss quadrature approximation has only minimal effect on those eigenvectors of H_0 that are almost orthogonal to the eigenvectors of Δ. In this context, only the eigenvectors that correspond to the infinite eigenvalues are meant when the eigenvectors of Δ are referred. $\mathbf{x}_1$, the only eigenvector of Δ_2 that has an infinite eigenvalue, was found to be nearly orthogonal to the lowest eigenvectors of H_0. The measure for the orthogonality is the inner product $\mathbf{v}_j^{\mathrm{T}}\mathbf{x}_i$, where $\mathbf{v}_j$ is the jth eigenvector of H_0 and $\mathbf{x}_i$ is one of the eigenvectors of the Δ matrix. Since $\mathbf{x}_1$ has already been handled, it is enough to show that the inner product is small for $\mathbf{x}_2$ and $\mathbf{x}_3$.

The inner product $\mathbf{v}_j^{\mathrm{T}}\mathbf{x}_1$ is approximated in Appendix A, and products $\mathbf{v}_j^{\mathrm{T}}\mathbf{x}_2$ and $\mathbf{v}_j^{\mathrm{T}}\mathbf{x}_3$ can be approximated using the same technique. As $\mathbf{x}_3$ is a linear combination of $\mathbf{x}_1$ and $\tilde{\mathbf{x}}_3$, it is enough to approximate the inner product $\mathbf{v}_j^{\mathrm{T}}\tilde{\mathbf{x}}_3$. Going through the first steps of the appendix for $\mathbf{v}_j^{\mathrm{T}}\mathbf{x}_2$ results in an equation that is equivalent to Eq. (A.4), namely

$$\mathbf{v}_j^{\mathrm{T}}\mathbf{x}_2 = \frac{1}{K_2}h^{-1/2}\sum_{i=1}^{N/2}\left[t_{2i-1}\Psi_j(t_{2i-1}) - t_{2i}\Psi_j(t_{2i})\right]. \tag{4.94}$$

A similar equation can be derived also for $\mathbf{v}_j^{\mathrm{T}}\tilde{\mathbf{x}}_3$, and it results in

$$\mathbf{v}_j^{\mathrm{T}}\tilde{\mathbf{x}}_3 = \frac{1}{\tilde{K}_3}h^{-3/2}\sum_{i=1}^{N/2}\left[t_{2i-1}^2\Psi_j(t_{2i-1}) - t_{2i}^2\Psi_j(t_{2i})\right]. \tag{4.95}$$

Here, t_i is the discrete position variable that defines the positions of the mesh points, i.e.

$$t_i = x_{a_i} = h a_i \,.$$

For $\mathbf{v}_j^{\mathrm{T}}\mathbf{x}_1$, the sum in the inner product was approximated by an integral of the derivative of the wave function. As the sum in the inner product $\mathbf{v}_j^{\mathrm{T}}\mathbf{x}_2$ has the wave functions multiplied with the corresponding position variables, but is otherwise similar to the sum in $\mathbf{v}_j^{\mathrm{T}}\mathbf{x}_1$, the sum can be approximated as

$$\begin{aligned}
\mathbf{v}_j^{\mathrm{T}}\mathbf{x}_2 &\approx -\frac{1}{2K_2} h^{-1/2} \int_{t_1}^{t_{N+1}} \frac{d}{dx}\big(x\Psi_j(x)\big)\, dx \\
&= -\frac{1}{2K_2} h^{-1/2} \big(a_{N+1}\Psi_j(t_{N+1}) - a_1\Psi_j(t_1)\big) \\
&= -\frac{1}{2K_2} \sqrt{h}\, \big(a_{N+1}\Psi_j(t_{N+1}) - a_1\Psi_j(t_1)\big)\,.
\end{aligned} \tag{4.96}$$

It is thus easy to generalize the bound for the absolute value of the inner product $\mathbf{v}_j^{\mathrm{T}}\mathbf{x}_1$ in Eq. (A.15) to cover $\mathbf{v}_j^{\mathrm{T}}\mathbf{x}_2$ by making the substitutions

$$K_1 \to K_2\,, \quad \Psi_j(t_i) \to a_i\Psi_j(t_i)\,.$$

Similarly, the inner product $\mathbf{v}_j^{\mathrm{T}}\tilde{\mathbf{x}}_3$ can be approximated as

$$\begin{aligned}
\mathbf{v}_j^{\mathrm{T}}\tilde{\mathbf{x}}_3 &\approx -\frac{1}{2\tilde{K}_3} h^{-3/2} \int_{t_1}^{t_{N+1}} \frac{d}{dx}\big(x^2\Psi_j(x)\big)\, dx \\
&= -\frac{1}{2\tilde{K}_3} \sqrt{h}\, \big(a_{N+1}^2\Psi_j(t_{N+1}) - a_1^2\Psi_j(t_1)\big)\,,
\end{aligned} \tag{4.97}$$

and the upper bound of the absolute value is the same as that of $\mathbf{v}_j^{\mathrm{T}}\mathbf{x}_1$ with the substitutions

$$K_1 \to \tilde{K}_3\,, \quad \Psi_j(t_i) \to a_i^2\Psi_j(t_i)\,.$$

The upper bounds for $|\mathbf{v}_j^{\mathrm{T}}\mathbf{x}_2|$ and $|\mathbf{v}_j^{\mathrm{T}}\tilde{\mathbf{x}}_3|$ are now known, and the final task is to show that they are small regardless of the small differences to the bound of $|\mathbf{v}_j^{\mathrm{T}}\mathbf{x}_1|$. It could potentially be an issue that the values of the wave functions are multiplied by a_1 and a_{N+1} or a_1^2 and a_{N+1}^2 because the absolute values of a_1 and a_{N+1} grow when the number of mesh points is increased, which is evident from Eq. (4.90). This problem, however, is fixed by considering the normalization coefficients K_2 and $\tilde{K}_3$. Very rough lower bounds for these coefficients are

$$K_2 = \sqrt{\sum_{i=1}^{N} a_i^2} \;\geq\; |a_1| = |a_N|$$

and

$$\tilde{K}_3 = \sqrt{\sum_{i=1}^{N} a_i^4} \;\geq\; a_1^2 = a_N^2\,.$$

Therefore, taking the normalization into account results in

$$\frac{1}{K_2}a_1 \leq \frac{a_1}{|a_1|} = -1\,, \quad \frac{1}{K_2}a_{N+1} \leq \frac{a_N+1}{|a_N|} = -1 + \frac{1}{|a_N|}$$

and

$$\frac{1}{\tilde{K}_3}a_1^2 \leq \frac{a_1^2}{a_1^2} = 1\,, \quad \frac{1}{\tilde{K}_3}a_{N+1}^2 \leq \frac{(a_N+1)^2}{a_N^2} = 1 + \frac{2}{a_N} + \frac{1}{a_N^2}\,.$$

The normalization thus neutralizes the effect of the multiplications with a_1 and a_{N+1} or a_1^2 and a_{N+1}^2 in the bounds of $|\mathbf{v}_j^{\mathrm{T}}\mathbf{x}_2|$ and $|\mathbf{v}_j^{\mathrm{T}}\tilde{\mathbf{x}}_3|$ respectively. Consequently, the same arguments that were used to conclude that the upper bound for $|\mathbf{v}_j^{\mathrm{T}}\mathbf{x}_1|$ is small can be used also for the bounds of $|\mathbf{v}_j^{\mathrm{T}}\mathbf{x}_2|$ and $|\mathbf{v}_j^{\mathrm{T}}\tilde{\mathbf{x}}_3|$. The absolute values of the inner products $\mathbf{v}_j^{\mathrm{T}}\mathbf{x}_2$ and $\mathbf{v}_j^{\mathrm{T}}\tilde{\mathbf{x}}_3$ are thus small and vanish at the limit $h \to 0$.

Based on the previous analysis, it can be concluded that those eigenvectors of the difference matrices Δ_4 and Δ_6 that correspond to the infinite eigenvalues are almost orthogonal to some of the lowest eigenvectors of the Hamiltonian matrix H_0, where the Gauss quadrature approximation has used. Therefore, the addition of Δ_4 or Δ_6 to H_0 has only minimal effect to those lowest eigenvectors, and the Gauss quadrature proves to be a very accurate approximation. Convergence of single-electron states in a 2-D fourth degree polynomial potential is tested in Section 5.1, and the results are indeed promising as the convergence is almost as fast as for the parabolic potential. Based on the success of the Gauss quadrature approximation for the polynomials of degree two, four and six, I present a hypothesis that it is an extremely accurate approximation for any polynomial potential, and it becomes exact at the limit $h \to 0$. The numerical results seem to agree with this hypothesis, but a rigorous proof is beyond the scope of this thesis.

Chapter 5

Computational results

5.1 Convergence tests and comparison

This section is dedicated to testing how well the Lagrange mesh method, combined with the exact diagonalization, performs in solving the single- and two-electron eigenstates in a quantum dot system. The convergence rates of the results are compared to a similar, but more established method, which uses Gaussian basis functions instead of the Lagrange sinc functions. This method has been used in the article by Nielsen et al. [12], which contains the derivation of the Hamiltonian matrix elements. The exact implementation is the same as in [35]. All the tests are are performed for systems without a global uniform magnetic field, i.e. $B = 0$.

Lagrange mesh method has two computational parameters that have nothing to do with the physical system. These are the size of the simulation area $L_x \times L_y$ and the number of mesh points $N_x \times N_y$. The mesh spacing $h = L_x/N_x = L_y/N_y$, i.e. the distance between the basis functions, is optimized so that the average relative error of the energies of 24 lowest eigenstates is minimized. In practice, this done by calculating the 24 lowest energies with a chosen range of different sizes of the simulation area, and these energies are compared to the accurate ones. Due to a restriction in the implementation of the program that was used to calculate the interaction matrix elements, all the many-particle results, and most of the single-particle ones also, are calculated using a mesh for which $N_x = N_y = N$.

The method that uses localized Gaussian basis functions, which will be called Gauss grid method from now on, has one extra parameter compared to the Lagrange mesh method, the Gaussian width a. Therefore, both the grid spacing h and the Gaussian width a must be optimized for each chosen set of $N_x \times N_y$ basis functions. Since the Gauss grid method is purely variational unlike the Lagrange mesh method, the smallest obtained energy is always the most accurate. This property is exploited by choosing h and a so that the weighted sum of the 24 first single-particle energies is minimized. The weights in the sum decrease towards the highest states because the lowest states are more likely occupied. The grid shape that was used for the single quantum dot was $N_x = N_y = N$, and for the double quantum dot it was chosen to be $N_x = 2N_y$.

The term relative error is used throughout this section, and it is defined as

$$\epsilon_r = |E - E_{ref}|/E_{ref}, \tag{5.1}$$

where the reference energy E_{ref} is either the exact energy, if that is available, or an energy that is as accurate as possible with one of the numerical methods. Choosing E_{ref} wisely is problematic in the case of piecewise parabolic potential because the exact solution is not available. Since the Gauss grid method seems to be better in terms of convergence and it can be considered as a generally accepted method, it is used to calculate the reference energy when the potential is not rounded. However, the Gauss grid method does not support the rounding method that is used for the Lagrange mesh potential, and therefore, the reference energies for the rounded potential have to be calculated using the Lagrange method.

The most time consuming part of the computation of many-electron states using the Lagrange mesh or Gauss grid method is the calculation of the interaction matrix elements. The computational effort to calculate these elements depends on the number of Lagrange or Gaussian basis functions as $O(N_{basis}^4)$ and on the size of the single-electron basis also as $O(N_s^4)$. Although the work is heavily parallelized using a GPU for the computation, it takes a couple of days to compute the interaction elements for 24 single-electron states using a basis size $N_{basis} = 400$. Since the computation time depends strongly on the basis size of the single-electron states, and so does the accuracy of the results, all the computational errors are presented as a function of it.

5.1.1 Single-electron states in one quantum dot

A quantum dot potential with one parabolic minimum is a good system to start the tests from because the single-electron states are analytically solvable in this case. The single-electron states in a parabolic potential are known as the Fock-Darwin states, and their energies are simply

$$E_{nl} = \omega(2n + |l| + 1)\,, \tag{5.2}$$

where ω is the confinement strength of the potential, $n \in \mathbb{N}$ and $l \in \mathbb{Z}$. The energies of the Fock-Darwin states offer the reference energies E_{ref} to which the numerically calculated energies can be compared.

The single-electron energies of the 24 lowest eigenstates in a parabolic quantum dot with the confinement strength $\hbar\omega = 4$ meV were computed using the Lagrange mesh and the Gauss grid methods, and the relative errors were calculated using the exact Fock-Darwin energies as reference values. The average of the relative errors as a function of the number of basis functions is shown in Fig. 5.1. The energies that were calculated using the Gauss grid method seem to converge substantially faster than the ones obtained with the Lagrange mesh method. The difference between the errors for each N_{basis} is about two orders of magnitude, and the Lagrange mesh method needs about twice the amount of basis functions compared to the Gauss grid method to reach the same accuracy. If the number of basis functions is doubled, the

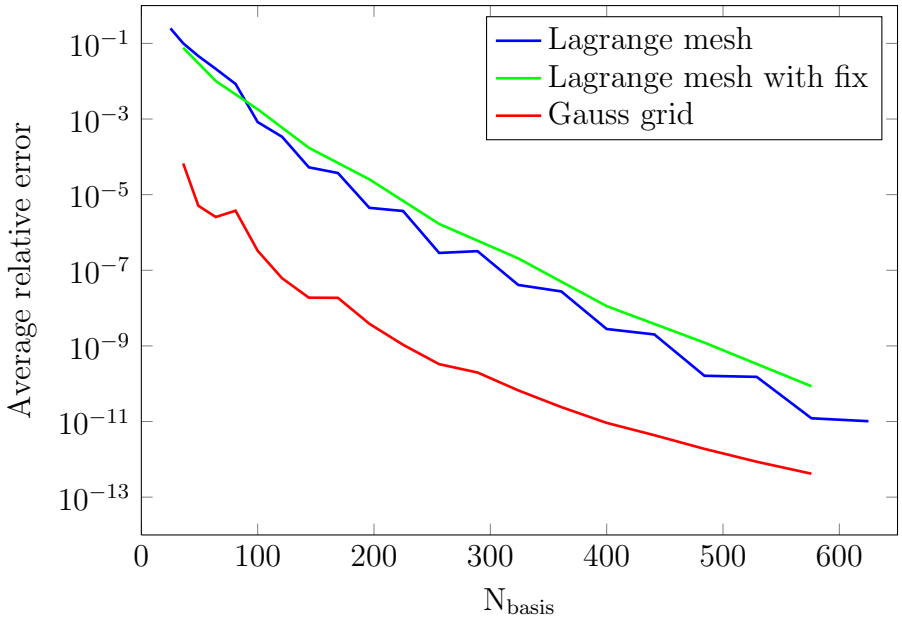

Figure 5.1: Convergence of the single-electron energies in one parabolic quantum dot as a function of the number of basis functions.

computation time increases by almost a factor of 16 because the computational effort to calculate the interaction elements behaves as $O(N_{basis}^4)$. The difference between the convergence speeds is thus quite dramatic. This comparison is somewhat unfair though, since the ground state of an electron in a parabolic quantum dot is exactly a Gaussian function, and consequently, the Gaussian functions are expected to be a good basis also for the higher states.

The green curve in Fig. 5.1 shows the convergence of the energies that have been calculated using the fix to get rid of the error caused by the Gauss quadrature approximation. These energies have been obtained by diagonalizing the H_{fix} Hamiltonian of Eq. (4.65). The convergence of the energies as a function of basis size is about as fast as those calculated without the fix. However, the point of the fix is to restore the variational principle, which is broken by the Gauss quadrature approximation. Since the Fig. 5.1 shows only the magnitude of the errors, it is impossible to say whether the energies are above or below the exact ones. For that purpose, the count of signs of the errors have been plotted in Fig. 5.2 for energies obtained without and with the Gauss quadrature fix. The red bars indicate that the sign of the energy difference $E - E_{ref}$ is positive, as it should be, and the blue bars show if some of the differences have been negative. The signs have been calculated separately for each of the 24 lowest states and then counted. The error sign counts for the energies that have been calculated without the fix show that for almost all values of N ($N_{basis} = N^2$), there are eigenstates that have energy lower than the exact one, and the variational principle is thus broken. The sign count bars for the energies that have been obtained using the fix are almost all red, meaning that the calculated energies have been higher than the exact ones. Only for $N = 24$ there have been a couple of states that have had lower energy than they should have, but

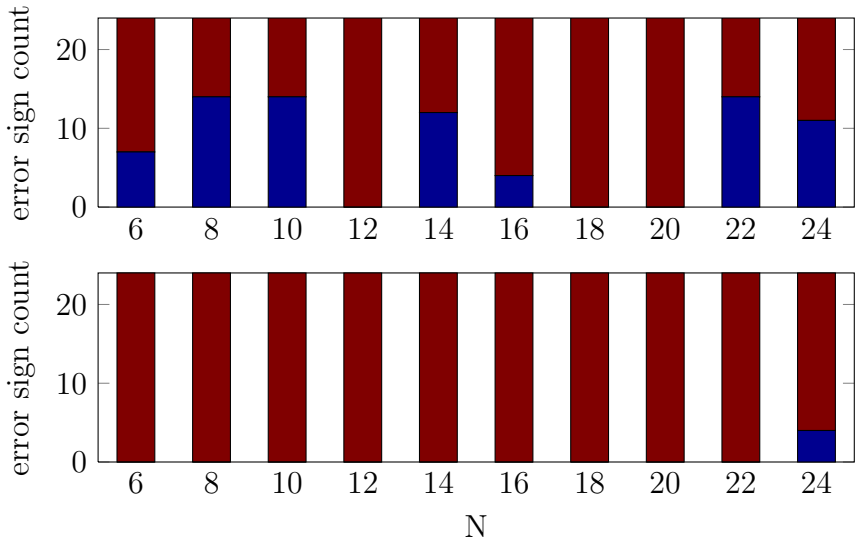

Figure 5.2: Count of signs (blue=-1, red=+1) of the errors of single-electron energies in a parabolic quantum dot. The energies have been calculated without the Gauss quadrature fix in the upper plot and with the fix in the lower plot. The signs of the errors have been calculated for the 24 lowest eigenstates. $N_{basis} = N^2$.

that is probably due to the fact that these energies have reached the numerical precision, and the lower energy is caused by a numerical error. The Gauss quadrature fix thus seems to work as it should, but unfortunately it can be used only for a potential with one parabolic minimum.

5.1.2 Single-electron states in double quantum dot

Since the single-particle states in a single parabolic quantum dot can be solved analytically, there is no point in calculating them numerically. The simplest quantum dot potential that is not solvable by analytical means is a double quantum dot with two parabolic minima. This is also a physically interesting system because it can be used to model the quantum dot based singlet-triplet qubits. The piecewise parabolic potential is constructed using the minimum function as in Eq. (2.4), and for the case of double quantum dot it becomes

$$V(\mathbf{r}) = \frac{1}{2}\omega^2 \min_{1,2} \left\{ |\mathbf{r} - \mathbf{R}_1|^2, |\mathbf{r} - \mathbf{R}_2|^2 \right\} , \tag{5.3}$$

where $\mathbf{R}_1$ and $\mathbf{R}_2$ are the locations of the minima, and the potential has been expressed in effective atomic units. In Lagrange mesh method, the simulation area is square shaped, unless otherwise stated, and the optimal locations of the minima are at the diagonal of the square. If the wanted distance between the minima is $2d$, the locations of the minima are $\mathbf{R}_1 = \frac{d}{\sqrt{2}}(\hat{\mathbf{e}}_x - \hat{\mathbf{e}}_y)$ and $\mathbf{R}_2 = -\frac{d}{\sqrt{2}}(\hat{\mathbf{e}}_x - \hat{\mathbf{e}}_y)$. The Gauss grid method handles only double quantum dot potentials where the minima are placed along any line that is parallel to either of the coordinate axes. The

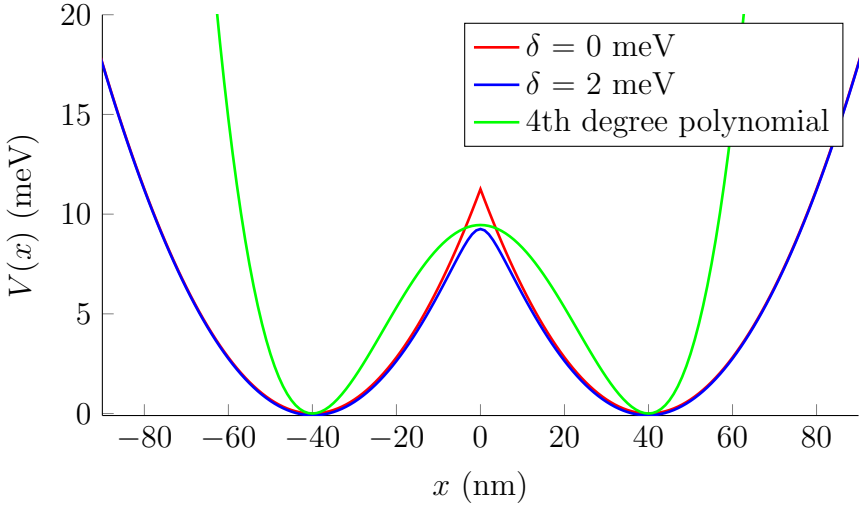

Figure 5.3: Cross-section of a double quantum dot potential where the minima are located on the x-axis at $x = -40$ nm and $x = 40$ nm. The red curve shows the non-rounded, i.e. $\delta = 0$, potential and the blue curve a potential rounded with off-diagonal elements $\delta = 2$ meV. The confinement strength of the potential is $\omega = 4$ meV. The green curve shows the 4th degree polynomial potential described in Eq. (5.4).

locations of the minima can thus be chosen to be along the x-axis at $\mathbf{R}_1 = d\hat{\mathbf{e}}_x$ and $\mathbf{R}_2 = -d\hat{\mathbf{e}}_x$.

A cross-section of a potential where the minima are placed along the x-axis at $x = -40$ nm and $x = 40$ nm is shown in Fig. 5.3. Since the minimum function chooses the parabolic minimum that gives a lower value at each point $\mathbf{r}$, it causes a point where the derivative of the potential is discontinuous. In this case it is located at $x = 0$. The sharp kink caused by the discontinuity of the derivative can be rounded easily using the following method. The values of each parabolic minimum at point $\mathbf{r}$ are placed on the diagonal of the square matrix V, and the off-diagonal elements are filled with a constant δ that defines the magnitude of the rounding. Then the lowest eigenvalue of V is chosen to be the value of the potential at point $\mathbf{r}$. In the case of a double quantum dot, V is a 2×2 matrix with diagonal elements $\mathrm{V}_{11} = \frac{1}{2}\omega^2|\mathbf{r}-\mathbf{R}_1|^2$ and $\mathrm{V}_{22} = \frac{1}{2}\omega^2|\mathbf{r}-\mathbf{R}_2|^2$, and with off-diagonal elements $\mathrm{V}_{12} = \mathrm{V}_{21} = \delta$. If the minima are located at $x = -d$ and $x = d$, the eigenvalues of V are

$$\lambda_{+/-} = \frac{1}{2}\omega^2(x^2 + y^2 + d^2) \pm \sqrt{\delta^2 + \omega^4 d^2 x^2}\,,$$

and the smaller eigenvalue is obviously λ_-. The cross-section of the rounded potential at $y = 0$ is shown in Fig. 5.3. $\delta = 2$ meV are used as the off-diagonal elements.

The convergence of the energies of single-electron states was tested for a double quantum dot potential where the minima are located at a distance of $2d = 80$ nm from each other and the confinement strengths of the parabolic minima were set to

$\hbar\omega = 4$ meV. The average relative errors of the 24 lowest eigenstates were calculated using the Lagrange mesh method with and without rounding and the Gauss grid method and they are shown in Fig. 5.4. The accurate reference energies for the non-rounded potential were computed using the Gauss grid method with large grid of $50 \times 25 = 1250$ Gaussian basis functions. The energies obtained with the Lagrange mesh method are comparable to the ones obtained with the Gauss grid method because the geometry of the minima locations should not affect the energies of the states. The Gauss grid method cannot handle the potential that is rounded using the off-diagonal matrix elements though, and thus the energies of the states in the rounded potential are compared to ones that are calculated using the Lagrange mesh method with a huge mesh of $70 \times 70 = 4900$ basis functions and a simulation area of size $L = 250$ nm.

Fig. 5.4 shows that Lagrange mesh method handles the piecewise parabolic double quantum dot potential even worse than the single dot potential when compared to the Gauss grid method. The average relative error of the single-electron energies seems to get stuck at a bit below 10^{-3}, whereas the average relative error of the energies obtained with the Gauss grid method reaches at least 10^{-6} and seems to decrease further if the number of basis functions would be increased. Even if a mesh size of 66×66, which corresponds to $N_{basis} = 4356$, is used for the Lagrange mesh method, the average relative difference is still 1.51×10^{-4}. The Lagrange mesh method handles the rounded potential much better, reaching average relative error in the order of 10^{-5} and it would still decrease if the basis size would be increased.

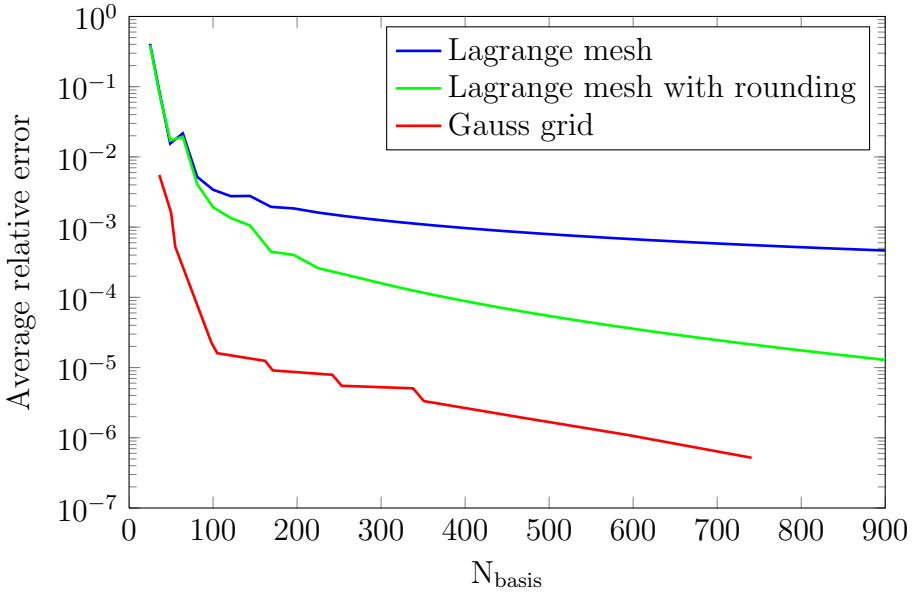

Figure 5.4: Average relative errors of the single-electron energies in a piecewise parabolic double quantum dot as a function of the number of basis functions. The distance between the minima is 80 nm and their confinement strength $\omega = 4$ meV. The off-diagonal elements that were used to round the intersection of the minima in the rounded version were chosen to be $\delta = 2$ meV.

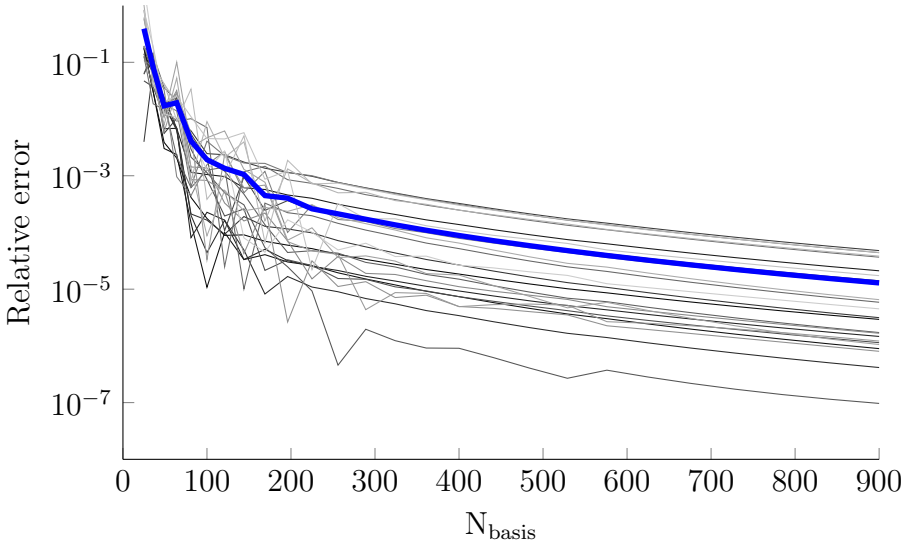

Figure 5.5: Convergence of the energies of each single-electron state in a rounded piecewise parabolic double quantum dot. The gray lines show the relative error of the energy of each state. The shades of gray have been assigned so that the lowest state has the darkest line and the lines lighten gradually towards the highest state. The thick blue line shows the average.

The reason why the piecewise potential without rounding is especially hard for the Lagrange mesh method is clear. The potential matrix elements that are obtained using the Gauss quadrature approximation seem to be extremely accurate only for a polynomial potential. This feature is discussed in Chapter 4. Therefore, the potential matrix elements are accurate for a polynomial that passes through all the mesh points. Since the derivative discontinuity of the potential can only be represented by an infinite degree polynomial, and it cannot be even approximated well by a polynomial of any reasonable degree, the potential matrix elements of the Lagrange mesh method are approximated badly. The elements for the rounded potential are approximated much better because the rounding makes at least the first derivative of the potential continuous. The potential matrix elements of the Gauss grid method, on the other hand, are calculated analytically taking the discontinuity of the potential explicitly into account.

Since the Lagrange mesh method works much better if the potential is rounded, all the physical results are calculated using the rounded potential. To show that none of 24 eigenstates converge much worse than they do in average, the relative errors of each of the states have been plotted in Fig. 5.5.

To strengthen the set of arguments which support the hypothesis that the potential matrix elements of Lagrange mesh are accurate for any polynomial potential, the double quantum dot potential was constructed from a fourth degree polynomial

$$V(x,y) = \alpha((x-d)^2(x+d)^2 + \beta y^4), \tag{5.4}$$

where $2d$ is the distance between the minima and α and β were tuned so that the resulting potential looked somewhat similar to the piecewise parabolic potential ($\alpha =$

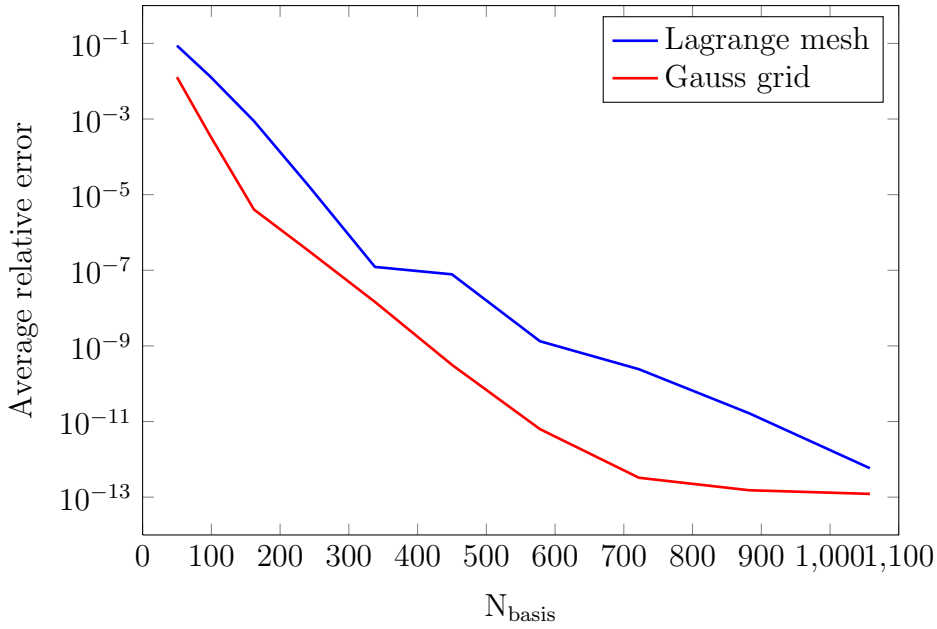

Figure 5.6: Average relative errors of the single-electron energies in a fourth degree polynomial double quantum dot as a function of the number of basis functions.

0.0033, $\beta = 8$). Fig. 5.6 shows the convergence of the single-electron state energies in the polynomial double quantum dot. Since the potential matrix elements for the Gauss grid method can be calculated analytically for any polynomial, the accurate reference energies were calculated using a Gauss grid of size 62×31. The energies were calculated also with the same grid sizes that were used with the Lagrange mesh for comparison. This is the only case where a rectangular Lagrange mesh, $N_x = 2N_y$, was used. The single-electron energies in the polynomial double quantum dot potential converge very fast using either of the two methods. The Lagrange mesh method actually performs a bit better in relation to the Gauss grid method when compared to the single quantum dot case. A polynomial potential, or potential that can be approximated well by a polynomial of a low degree, seems to be the key to a fast convergence of the Lagrange mesh results. Since the rounded piecewise polynomial potential is far more flexible than the pure polynomial potential and it seems to offer reasonably good results, it is used in the computations of the physical results.

5.1.3 Two-electron states in double quantum dot

To study the control of singlet-triplet qubits, one has to calculate the two-electron states in the double quantum dot. The convergence of the two-electron states depends largely on the convergence of the single-particle states, but also on the quality of the interaction matrix elements. Most of the work to calculate the interaction matrix elements for the single-particle states that are computed using the Lagrange mesh method can be done analytically. Some numerical integrals are also needed,

but since the integrals are independent of the simulated system, they need to be calculated only once, and thus it is no problem to calculate them up to the numerical precision. The interaction matrix elements of the states that are calculated using the Gauss grid method, on the other hand, can be calculated completely analytically. Therefore, the accuracy of the interaction elements depends only on the single-electron wave functions in both methods.

The convergence of the eigenvalues and the eigenvectors of the Hamiltonian matrix, i.e. the energies and the wave functions, are obviously correlated, and thus the accuracy of the two-electron states can be predicted from the convergence status of the single-electron energies. However, the average relative error of the 24 lowest single-electron energies and the relative errors of the lowest two-electron states can differ significantly, because the two electrons are likely to occupy only a few of the lowest single-electron states. Therefore, the convergence of the lowest singlet and triplet states was checked by calculating their relative errors. The relative errors of the singlet and triplet states that were calculated using the exact diagonalization method from the Lagrange mesh and Gauss grid single-electron states are shown in Fig. 5.7. The potential is non-rounded in both cases and the accurate reference energies were calculated from the single-electron states that were obtained using the Gauss grid method with grid size 34×17, which corresponds to $N_{basis} = 578$. The difference between the two methods seems to be even more dramatic than what it was in the comparison of the average relative error of the single-electron energies. The relative errors of the Lagrange mesh singlet and triplet states in a non-rounded potential are not that bad though, as for $N_{basis} = 400$ the errors go a bit below 10^{-4}, which should be more than enough for qualitatively correct physical results. On the other hand, the errors of the Gauss grid singlet and triplet states are remarkably small, reaching below 10^{-7} when $N_{basis} = 450$.

Although the Lagrange mesh method seems to give reasonably accurate two-electron states in a non-rounded piecewise parabolic potential, the single-electron states in the rounded potential were found to be considerably more accurate than in the non-rounded potential. Therefore, it is safe presume that the two-electron states should also converge faster if the rounded potential is used. The convergence of the singlet and triplet states in the rounded potential is somewhat harder to study than the convergence of those in the non-rounded potential because the reference energies cannot be calculated using the Gauss grid method. The energies can only be compared to those that are calculated using as large mesh as possible. Since the current implementation limits the largest possible mesh size for the calculation of the interaction matrix elements to 20×20, that mesh was used to calculate the reference energies for the rounded and non-rounded potential. The relative errors of the lowest singlet and triplet states in both potentials are shown in Fig. 5.8. The results are as expected, the two-electron energies converge faster in the case of the rounded potential.

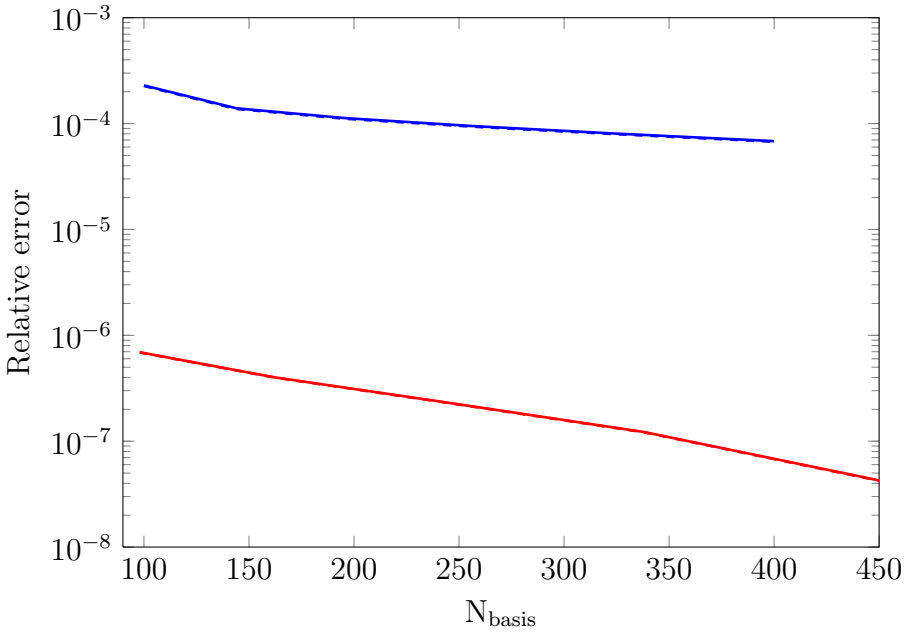

Figure 5.7: Convergence of the energies of singlet (solid line) and triplet (dashed line) states ($|S\rangle$ and $|T_0\rangle$) in a piecewise parabolic double quantum dot without rounding of the potential. The blue curves show the errors of the Lagrange mesh method and the red ones the errors of the Gauss grid method. The errors of singlet and triplet states are so close together that the difference can barely be seen.

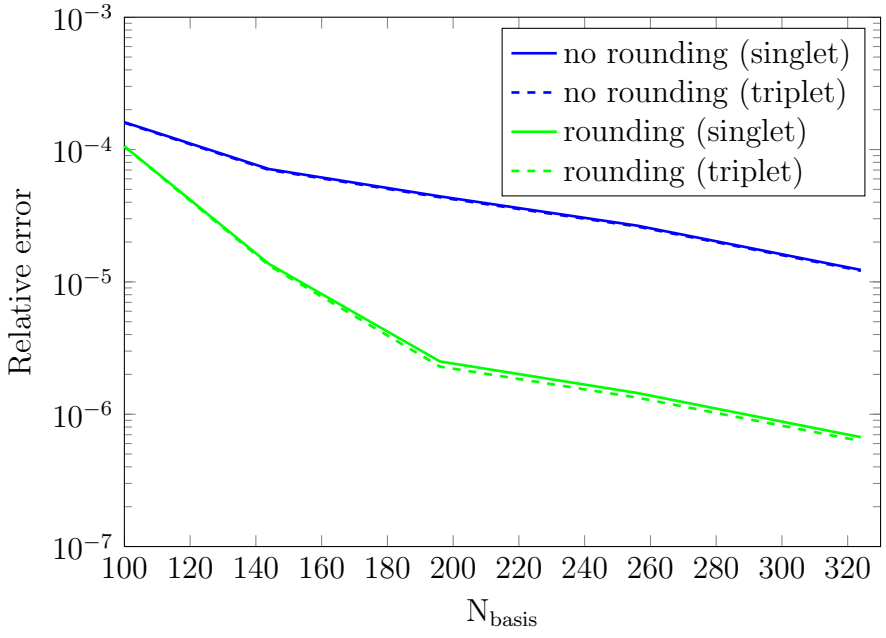

Figure 5.8: Convergence of the energies of singlet and triplet states ($|S\rangle$ and $|T_0\rangle$) in a piecewise parabolic double quantum dot with rounding compared to potential without rounding.

5.2 Control of qubit

5.2.1 Implementation of tunable detuning and magnetic field gradient

As was discussed in Section 2.3, complete control of a qubit requires a way to carry out rotations of the state around two axes of the Bloch sphere. In the case of singlet-triplet qubits, the rotations around the z-axis are driven by the exchange splitting J between the basis states of the qubit, $|S\rangle$ and $|T\rangle$, and the rotations around the x-axis are driven by a magnetic field gradient between the two quantum dots. The exchange splitting can be controlled by changing the detuning between the dots, i.e. by shifting the energy levels of the parabolic minima. To simulate the rotations, one must implement a way to control the detuning and the magnetic field gradient efficiently during the time evolution of the state.

When considering the effect of the magnetic field gradient, the corresponding vector potential is omitted and only the Zeeman term is taken into account. This approximation is reasonable because difference of the magnetic fields between the dots is typically quite small and the magnetic fields can be chosen to be parallel to the plane of the dots, which minimizes the effect on the single-electron orbitals [36]. The detuning and the energy separation due to the Zeeman effect can be added to the symmetric base potential by shifting the levels of the two parabolic minima corresponding to the amount of detuning ε or magnetic field gradient ΔB. The detuning is defined so that a positive value of ε symmetrically raises the left potential minimum and lowers the right one. The left minimum is thus shifted by $+\varepsilon/2$ and the right one by $-\varepsilon/2$. The rounded potential is obtained by adding the shifts due to the detuning to the diagonal elements of the rounding matrix V, i.e. $\mathrm{V}_{11} = \frac{1}{2}\omega^2|\mathbf{r} - \mathbf{R}_1|^2 + \frac{\varepsilon}{2}$ and $\mathrm{V}_{22} = \frac{1}{2}\omega^2|\mathbf{r} - \mathbf{R}_2|^2 - \frac{\varepsilon}{2}$. The shifts due to the Zeeman energy difference are $+\frac{1}{2}g^*\mu_B\Delta B m_S$ for the left dot and $-\frac{1}{2}g^*\mu_B\Delta B m_S$ for the right one. Here, $m_S = \pm\frac{1}{2}$ depending on the z-component of the electron spin. A potential with a detuning of $\varepsilon = 4$ meV is shown in Fig. 5.9 together with the symmetric potential.

The effects of the detuning and the magnetic field gradient can added as corrections to the single-electron Hamiltonian

$$\tilde{H}_0 = H_0 + V_d(\mathbf{r}, \varepsilon) + 2V_Z(\mathbf{r}, \Delta B)S_z\,, \tag{5.5}$$

where

$$H_0 = -\frac{1}{2}\nabla^2 + V(\mathbf{r})\,.$$

The corrections $V_d(\mathbf{r}, \varepsilon)$ and $V_Z(\mathbf{r}, \Delta B)$ are calculated as the difference between the potential where the minima have been shifted due to detuning or magnetic field gradient and the symmetric base potential. $V_Z(\mathbf{r}, \Delta B)$ is calculated for an electron with $m_S = +\frac{1}{2}$ and the S_z operator takes the electron spin into account, hence the correction coefficient 2 has been added. The difference $V_d(\mathbf{r}, \varepsilon)$ between the base potential with $\varepsilon = 0$ and $\Delta B = 0$ and a potential with a detuning of $\varepsilon = 4$ meV is shown in Fig. 5.9. The difference $V_Z(\mathbf{r}, \Delta B)$ has exactly the same form. Changing

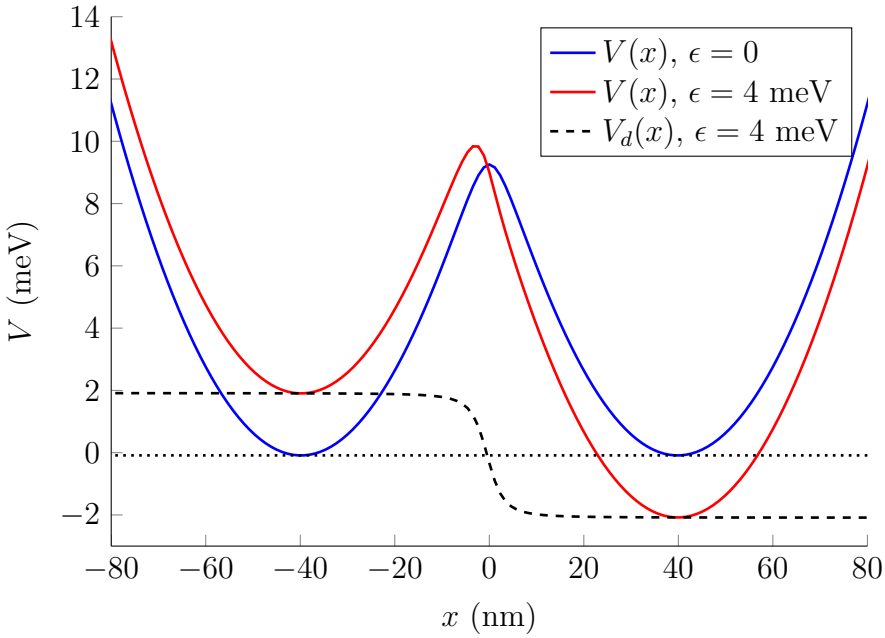

Figure 5.9: A non-detuned double quantum dot potential and a potential with a symmetric detuning of $\varepsilon = 4$ meV. V_d is the difference between the non-detuned and detuned potentials.

the detuning ε or the field gradient ΔB alters the single-electron states, and thus the states and the interaction matrix elements would have to be calculated for all values of ε and ΔB that are used in the dynamics. Therefore, it is not sensible to solve the single-electron states corresponding to this modified Hamiltonian. Instead, the single-electron states are calculated for the base Hamiltonian H_0 and the corrections are added to the second quantization Hamiltonian, which can be written as

$$H = \sum_{j,\sigma} E_j \, a^\dagger_{j\sigma} a_{j\sigma} + \sum_{i,j,k,l,\sigma} U_{ijkl} \, a^\dagger_{i\sigma} a^\dagger_{j\sigma} a_{k\sigma} a_{l\sigma} + \sum_{i,j,\sigma} (\varepsilon_c V_{d,ij} + \Delta B_c V_{Z,ij\sigma}) a^\dagger_{i\sigma} a_{j\sigma} \, . \quad (5.6)$$

σ denotes the spin z-component of the electron, i.e. $\sigma = m_S = \pm\frac{1}{2}$. The matrix elements of the corrections are

$$V_{d,ij} = \langle \psi_i | \, V_d(\mathbf{r}, \varepsilon_0) \, | \psi_j \rangle$$

and

$$V_{Z,ij\sigma} = \langle \psi_{i,\sigma} | \, 2V_Z(\mathbf{r}, \Delta B_0) S_z \, | \psi_{j,\sigma} \rangle = 2\sigma \langle \psi_{i,\sigma} | \, V_Z(\mathbf{r}, \Delta B_0) \, | \psi_{j,\sigma} \rangle \, .$$

As the single-electron states are constructed from the basis of Lagrange functions, the integrals of the elements can be approximated using the Gauss quadrature for the integrals of the Lagrange functions.

The matrix elements for the corrections due to detuning and magnetic field gradient are calculated using some convenient values of ε_0 and ΔB_0, such as $\varepsilon_0 = 1$ meV and $\Delta B_0 = 1$ T, and the relative values ε_c and ΔB_c are used to control the

detuning and the field gradient in the dynamics, i.e.

$$\varepsilon = \varepsilon_c \varepsilon_0, \quad \Delta B = \Delta B_c \Delta B_0 \,.$$

This approach is sensible because $V_d(\mathbf{r}, \varepsilon)$ and $V_Z(\mathbf{r}, \Delta B)$ depend approximately linearly on ε and ΔB_0 respectively, and thus

$$V_d(\mathbf{r}, \varepsilon) \approx \varepsilon_c V_d(\mathbf{r}, \varepsilon_0), \quad V_Z(\mathbf{r}, \Delta B) \approx \Delta B_c V_Z(\mathbf{r}, \Delta B_0) \,.$$

5.2.2 Controlling exchange splitting with detuning

The exchange splitting J, which drives the rotations of the state vector around the z-axis in the Bloch sphere, can be tuned by controlling the detuning ε between the quantum dots. To have exact control of the qubit rotations, one must know how J depends on the detuning. The energies of the three lowest two-electron states as functions of the detuning are shown in Fig. 5.10. The qubit basis states $|(1,1)S\rangle$ and $|(1,1)T_0\rangle$ are well separated from the third lowest state $|(2,0)S\rangle$ until the anti-crossing point of $|(1,1)S\rangle$ and $|(2,0)S\rangle$ at around $\varepsilon = 4.7$ meV. $(1,1)$ and $(2,0)$ denote the possible charge states of the qubit, i.e. one electron in each quantum dot or both electrons in the same dot respectively. To avoid transition to the higher singlet state, the detuning should be chosen somewhat below the anti-crossing point. The exchange splitting, defined as $J = E_{T_0} - E_S$, is shown in Fig. 5.11.

5.2.3 Qubit rotation tests

An effective Hamiltonian will now be created for the singlet-triplet qubit to have a theoretical basis for the rotations. Coish and Loss have derived the Hamiltonian for the two-electron states $|S\rangle$, $|T_0\rangle$, $|T_-\rangle$ and $|T_+\rangle$ [37]. In the exact diagonalization calculations, the many-body basis is constructed from the subspace of two-electron states with $\langle S_z \rangle = 0$. Therefore, $|T_-\rangle$ and $|T_+\rangle$ cannot be reached. In real devices, $|T_+\rangle$ and $|T_-\rangle$ are separated from the qubit basis state $|T_0\rangle$ by applying a strong uniform magnetic field B to the same direction as the magnetic field gradient between the quantum dots. The magnetic field produces a Zeeman splitting of $\pm g^* \mu_B B$ for the $|T_+\rangle$ and $|T_-\rangle$ states because those states have $m_S = \pm 1$. As the lowest $|S\rangle$ and $|T_0\rangle$ states are also energetically separated from the higher ones, which is evident from Fig. 5.10, the effective Hamiltonian can be constructed for the qubit basis states $|S\rangle$ and $|T_0\rangle$ only. The effective Hamiltonian in the basis $|S\rangle$ and $|T_0\rangle$ can be written as [37]

$$\mathrm{H} = \begin{bmatrix} 0 & \delta h/2 \\ \delta h/2 & J \end{bmatrix}, \tag{5.7}$$

where J is the exchange splitting and

$$\delta h = g^* \mu_B \Delta B.$$

The qubit basis states $|S\rangle$ and $|T_0\rangle$ are the eigenstates of the effective Hamiltonian when $\delta h = 0$. In this case, the time evolution operator is

$$e^{-it\mathrm{H}} = \begin{bmatrix} 0 & 0 \\ 0 & e^{-itJ} \end{bmatrix},$$

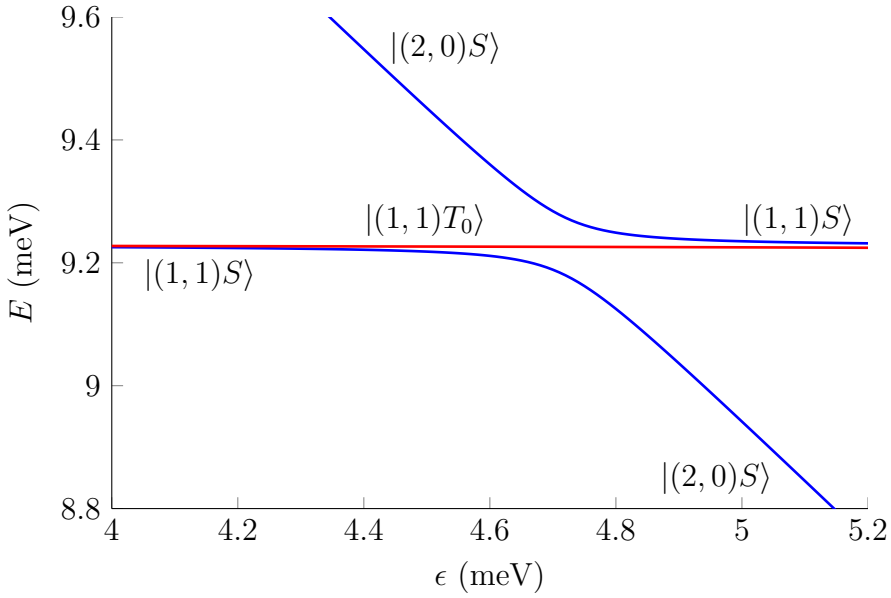

Figure 5.10: The energies of the lowest two-electron states as a function of detuning.

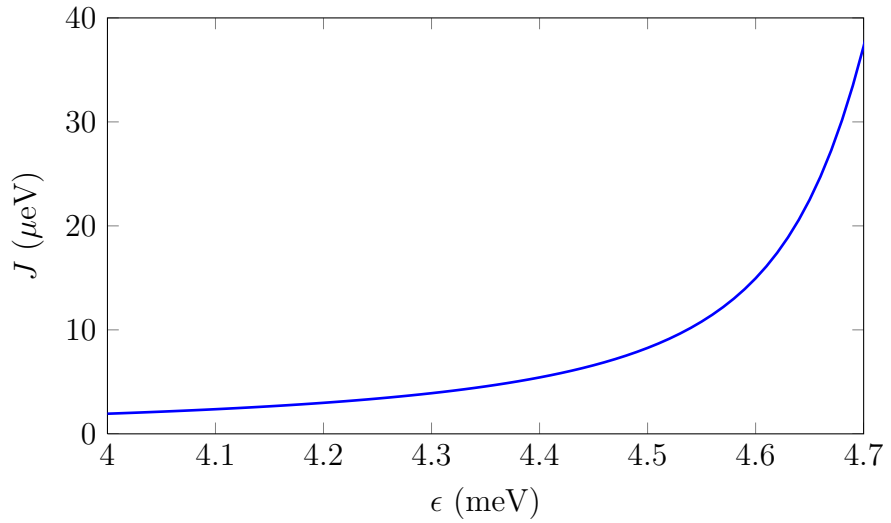

Figure 5.11: The exchange splitting J between the qubit basis states $|S\rangle$ and $|T_0\rangle$.

and the time evolution of a state vector corresponding to the qubit state $|\psi\rangle = \alpha|S\rangle + \beta|T_0\rangle$, i.e. $[\alpha\ \beta]^{\mathrm{T}}$, is simply

$$e^{-it\mathrm{H}}\begin{bmatrix}\alpha\\ \beta\end{bmatrix} = \begin{bmatrix}\alpha\\ e^{-itJ}\beta\end{bmatrix}.$$

For a singlet-triplet qubit with an exchange splitting J and no magnetic field gradient, the time evolution operator thus changes the complex phase of the triplet part of the state by amount $-\Delta tJ$. Recalling that $\beta = e^{i\phi}\sin(\frac{\theta}{2})$ in the Bloch sphere representation, the addition of the complex phase can be written as

$$e^{-i\Delta tJ}\beta = e^{i(\phi - tJ)}\sin(\frac{\theta}{2}),$$

i.e. the time evolution rotates the state vector in the Bloch sphere around the z-axis by angle $\phi_{rot} = -tJ$. In SI units, the rotation angle is given by $\phi_{rot} = -tJ/\hbar$.

If $\delta h \neq 0$, the singlet basis states are no longer the eigenstates of the Hamiltonian. Diagonalizing the Hamiltonian in the case $\delta h \neq 0$ and $J = 0$ gives the eigenvectors

$$\mathbf{u}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}, \quad \mathbf{u}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix},$$

which correspond to the eigenvalues $\lambda_1 = \delta h/2$ and $\lambda_2 = -\delta h/2$. The time evolution operator can now be decomposed into

$$e^{-it\mathrm{H}} = \mathrm{U}e^{-it\Lambda}\mathrm{U}^*$$

where $\mathrm{U} = [\mathbf{u}_1\ \mathbf{u}_2]$ and $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2)$. The operator $e^{-it\mathrm{H}}$ can be interpreted in the following way. First, U^* transforms the state vector to the eigenbasis of H. Then, the complex phases of the eigenstates are evolved by $e^{-it\Lambda}$. Finally, U transforms the state vector back to the basis of $|S\rangle$ and $|T_0\rangle$. $e^{-it\Lambda}$ adds a complex phase difference of $t(\lambda_2 - \lambda_1)$ between the eigenstates of the Hamiltonian. In the Bloch sphere representation this corresponds to a rotation of angle $\chi = t(\lambda_2 - \lambda_1)$ around the axis defined by the eigenvectors $\mathbf{u}_1$ and $\mathbf{u}_2$. In the case of $\delta h \neq 0$ and $J = 0$, $\mathbf{u}_1$ and $\mathbf{u}_2$ correspond to the states $|\uparrow\downarrow\rangle$ and $|\downarrow\uparrow\rangle$ respectively, and thus the rotations are around the x-axis of the Bloch sphere.

In a more general case of $\delta h \neq 0$ and $J \neq 0$, the eigenstates of the effective Hamiltonian are more complex, but they can be shown to correspond to a rotation axis

$$\mathbf{n} = J\hat{\mathbf{e}}_{\mathbf{z}} + \delta h\hat{\mathbf{e}}_{\mathbf{x}}, \tag{5.8}$$

which is tilted by an angle

$$\theta = \arctan\left(\frac{\delta h}{J}\right)$$

from the z-axis. The rotation angle obtained after time t is

$$\chi = t\sqrt{J^2 + \delta h^2}/\hbar.$$

Now that the theoretical basis for the rotations is known, the rotations can be simulated using the exact diagonalization based dynamics and compared with the theoretical results. The qubit that is studied has a rounded potential with two parabolic minima with a center to center distance of 80 nm and a confinement strength of $\hbar\omega = 4$ meV. The system is thus the same as the rounded double quantum dot that was used in the convergence tests. The Bloch sphere angles θ and ϕ are obtained from the qubit state $|\psi\rangle = \alpha|S\rangle + \beta|T_0\rangle$ using the two-electron S^2 operator, which gives

$$S^2|\psi\rangle = \alpha S^2|S\rangle + \beta S^2|T_0\rangle = 2\beta|T_0\rangle \Rightarrow \beta = \frac{1}{2}\langle T_0|S^2|\psi\rangle.$$

The previous formula is based on the fact that $|S\rangle$ and $|T_0\rangle$ are eigenstates of the S^2 operator with the corresponding quantum numbers $s = 0$ and $s = 1$, and the eigenvalue equation of S^2 is $S^2|s\rangle = s(s-1)|s\rangle$. In the Bloch sphere representation, $\beta = e^{i\phi}\sin(\theta/2)$, and thus θ can be extracted from $|\beta|^2 = \sin^2(\theta/2)$ and it can be used to obtain ϕ.

Simple rotations around the x- and z-axes provide a good starting point for the simulations. As the exchange splitting cannot be turned completely off, at $\varepsilon = 0$ the exchange splitting is $J = 0.189$ μeV, the magnetic field gradient ΔB must be large enough to ensure $\delta h \gg J$ which provides accurate rotations around the x-axis. It has been experimentally shown that $\Delta B \approx 200$ mT can be obtained using the dynamic nuclear polarization technique [22]. That is a bit too small to provide an accurate x-rotation for the qubit that is simulated in this thesis, and thus $\Delta B = 300$ mT, which is not too far from the experimentally obtained value, is used instead. This value of the magnetic field gradient corresponds to $\delta h = -7.641$ μeV. The rotation around the z-axis is done using a detuning of $\varepsilon = 4.5$ meV, which corresponds to $J = 8.264$ μeV. The rotation times are set to obtain rotations of an angle π by calculating the needed times from the theoretical formulas $t_x = \hbar\chi/\delta h$ and $t_z = \hbar\chi/J$, setting $\chi = \pi$. The simulated π rotations are shown in Fig. 5.12, and they seem to agree well with the theoretical considerations as the used rotation times produce almost perfect π rotations.

A more complicated rotation is obtained when $J \neq 0$ and $\delta h \neq 0$. In this case, the theoretical effective Hamiltonian predicts that the rotation axis is tilted by an angle $\theta = \arctan(\delta h/J)$ relative to the z-axis. The control of a qubit is usually based on two working points with fixed values of J and δh. In real double quantum dot devices, the magnetic field gradient is always present during the qubit rotations and it can be considered to be constant, if fluctuations of the hyperfine field are ignored. Since the magnetic field gradient cannot be turned off, the other working point must have rotations that are tilted by the angle θ. Ramon [38] has provided a pulse sequence that produces arbitrary rotations around the z-axis with working points $J \gtrsim \delta h$ and $J = 0$. A rotation with $J = 3.909$ μeV ($\varepsilon = 4.3$ meV) and $\delta h = 3.820$ μeV ($\Delta B = 150$ mT), which satisfies the condition $J \gtrsim \delta h$ for the other working point, is shown in Fig. 5.12. The duration of the rotation pulse has been adjusted to obtain a rotation angle of 2π and it was calculated from $t = \hbar\chi/\sqrt{J^2 + \delta h^2}$ with $\chi = 2\pi$. The obtained rotation angle is very close 2π, and the calculated rotation axis agrees well with the simulated rotation.

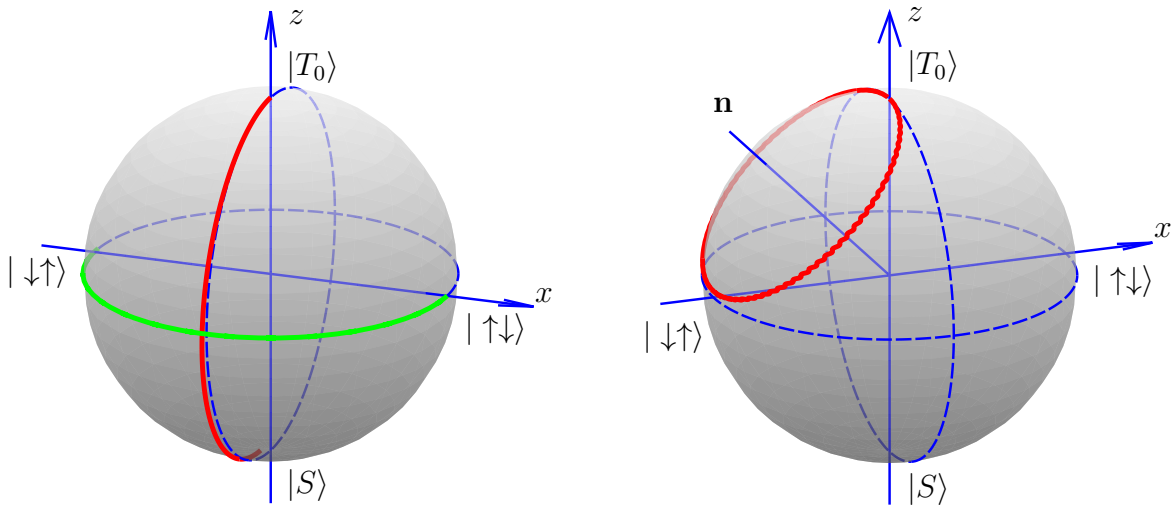

Figure 5.12: Simulated rotations of the state in the Bloch sphere. In the left plot, rotations of the state vector by an angle π around the x and z-axes are shown. The red curve shows a rotation around the x-axis using a magnetic field gradient of $\Delta B = 300$ mT, and the green curve shows a rotation around the z-axis using a exchange splitting $J = 8.264$ μeV ($\varepsilon = 4.5$ meV). In the right plot, the state is rotated by angle 2π around a combined axis $\mathbf{n} = J\hat{\mathbf{e}}_\mathbf{z} + \delta h\hat{\mathbf{e}}_\mathbf{x}$, where $J = 3.909$ μeV and $\delta h = -3.820$ μeV.

Chapter 6

Conclusions

An exact simulation of a quantum mechanical system on a classical computer is a difficult problem, as Feynman realized already in 1982. He could not anticipate the immense speed-up of the computers from those days, however, which has made simulations of certain simple systems possible. A two-dimensional system of a few electrons in an external confining potential is one of those, and one possible set of computational methods to study such a system was presented in this thesis.

The set of computational methods that was used in this work consists essentially of three parts; the Lagrange mesh method, the exact diagonalization and dynamics. Unlike exact diagonalization, which is used widely in computational physics, Lagrange mesh method has not achieved similar foothold despite the fact that it was properly introduced already in 1986 by Baye and Heenen [39]. One major problem with the Lagrange mesh method is that it is still not known why the method works as well as it does despite the fact that the Gauss quadrature, the main approximation used in the method, is generally bad for the individual potential matrix elements [40]. Therefore, the main goal of this work was to analyse the effect of the Gauss quadrature approximation on the eigenvalues and eigenstates of the Hamiltonian.

The greatest accomplishments of this work are the rigorous analysis of the Gauss quadrature approximation for a parabolic potential, the attempt to generalize the results of the analysis to any polynomial potential, and the comprehensive benchmark of the Lagrange mesh method. The mathematical analysis of the effect of the Gauss quadrature approximation in the case of a single parabolic potential showed that the error it causes decreases as the mesh spacing is decreased. According to the numerical convergence tests, the error caused by the Gauss quadrature is extremely small compared to the error caused by the discretization, but the variational principle is lost. A way to neutralize the error due to Gauss quadrature was found, however, and it can be used the restore the variational principle, if needed.

The effect of the Gauss quadrature was considered also for 4th and 6th degree polynomials. The analysis was not as rigorous as in the case of the 2nd degree polynomial, but it seems that the Gauss quadrature is a good approximation for those polynomials as well. Based on this result, a hypothesis that the Gauss quadrature approximation works well for a polynomial of arbitrary degree was made. If the Gauss quadrature approximation really is an accurate approximation for any

polynomial, it would imply that the Lagrange mesh method effectively makes a polynomial fit for the given potential so that it passes through the mesh points. Consequently, it would guarantee that the potential matrix elements obtained with the Gauss quadrature are good for any potential that can be approximated well by a polynomial. This could be the reason why the single-electron states converge relatively slow in the case of a piecewise parabolic potential, which has a derivative discontinuity.

The 24 lowest single-electron eigenstates in a single parabolic quantum dot and a double quantum dot consisting of two parabolic minima were computed for various mesh sizes using the Lagrange mesh and a more established method, which is called Gauss grid method in this thesis. For the double quantum dot, also the two lowest two-electron states were calculated. The convergence of the energies corresponding to these states was compared to find out which method is more efficient. The Gauss grid method seems to perform better in all aspects; both the single-electron states and the two-electron states converge faster for all potentials that were tested. Especially, the Gauss grid handles the piecewise parabolic potential much better than the Lagrange mesh. The Lagrange mesh method seems to handle the rounded double quantum dot potential reasonably well though, but the convergence in the rounded potential cannot be compared directly to the Gauss grid method because it can handle only piecewise polynomial potentials.

Although the states calculated using the Gauss grid method were found to converge faster in all potentials that were tried, the Lagrange mesh method is more flexible in many ways. The Gauss grid method handles only piecewise polynomial potentials and the geometry of the quantum dot locations is restricted. Furthermore, as the Gaussian basis functions are not orthogonal, the eigenvalue problem is generalized. The Lagrange mesh method, on the other hand, handles arbitrary potentials if enough mesh points are used, and the basis functions are orthonormal.

The Lagrange mesh single-electron states in the rounded potential were found to be accurate enough for the simulations of the qubit state control. The simulated qubit rotations were extremely close to the ones predicted by the theoretical considerations. The used values of detuning were close to the limit after which the accurate dynamics breaks, however. This is arguably caused by the fact that the single-particle states are calculated for a symmetric potential, which produces states that are either symmetric or antisymmetric. These states are far away from the highly asymmetric states that are obtained with a high detuning. This problem could perhaps be solved by calculating the states for a slightly asymmetric potential.

In summary, this thesis provides an explanation for the high accuracy of the Lagrange mesh method despite the seemingly crude Gauss quadrature approximation. It seems that the error due to the Gauss quadrature is small for any potential that can be approximated reasonably well by a polynomial with degree equal to the number of mesh points per dimension. The qubit rotation tests showed that the combination of the Lagrange mesh method and the exact diagonalization can be used to obtain highly accurate dynamics of the qubit state under moderate range of modifications to the base potential. As these modifications can be time dependent and they are easy to apply, the presented methods suit well for computational

studies of qubit state control. In addition to single qubits, these methods can be applied to systems of two qubits [35]. By taking the fluctuations of the hyperfine field into account as has been done by Särkkä [41], also the decoherence effects could be studied.

Appendix A

Approximating inner product $\mathbf{v}_j^{\mathrm{T}}\mathbf{x}$

The measure for the orthogonality of the eigenvectors of H_0 and the vector $\mathbf{x}$ is given by the inner product $\mathbf{v}_j^{\mathrm{T}}\mathbf{x}$, where $\mathbf{v}_j$ is the jth eigenvector of H_0. A series approximation for the inner product is derived in this appendix.

As the vector $\mathbf{x}$ is given by Eq. (4.13), the inner product can be written as

$$\mathbf{v}_j^{\mathrm{T}}\mathbf{x} = \frac{1}{\sqrt{N}}(v_1 - v_2 + v_3 - v_4 + \cdots + (-1)^{N-1}v_N), \tag{A.1}$$

where v_i are the components of the vector $\mathbf{v}_j$ and they correspond to the coefficients $c_{\mathbf{a}}$ in the wave function expansion, Eq. (3.10), so that

$$v_i = c_{-\frac{N+1}{2}+i}, \quad i \in [1, N].$$

Using the Lagrange conditions in Eq. (3.2), the wave function expansion can be written as

$$\Psi_j(x_a) = \sum_{a'} c_{a'}\lambda_{a'}^{-1/2}\delta_{a'a} = \frac{c_a}{\sqrt{h}}. \tag{A.2}$$

Here, a fact that the weights for the 1-D Lagrange mesh are $\lambda_a = h$, is used. Using the equations above to solve the vector components v_i gives

$$v_i = \sqrt{h}\Psi_j(x_{-\frac{N+1}{2}+i}) = \sqrt{h}\Psi_j(-\frac{L}{2} - \frac{h}{2} + hi). \tag{A.3}$$

Thus, $\frac{v_i}{\sqrt{h}}$ are the discretized values of the wave function $\Psi_j(x)$. As $\Psi_j(x)$ consists of a finite sum of infinitely differentiable functions, the Lagrange functions $L_a(x)$, also $\Psi_j(x)$ is infinitely differentiable. Therefore, the inner product $\mathbf{v}_j^{\mathrm{T}}\mathbf{x}$ can be related to an integral of the derivative of $\Psi_j(x)$.

By substituting the derived relation for the eigenvector components, Eq. (A.3), to Eq. (A.1), the inner product can be written as

$$\mathbf{v}_j^{\mathrm{T}}\mathbf{x} = \frac{1}{\sqrt{N}}\sqrt{h}\sum_{i=1}^{N/2}[\Psi_j(t_{2i-1}) - \Psi_j(t_{2i})], \tag{A.4}$$

where the position variable has been substituted with

$$t_i = x_{-\frac{N+1}{2}+i}$$

for better readability. It is assumed that N is even because otherwise $\Psi_j(t_N)$ is not included in the sum. N can be odd, however, because $\Psi_j(t_N)$ should be approximately zero. Since the position is discretized, it is possible to write $t_{2i} = t_{2i-1} + h$. Now, the terms in the sum can be written in the form of difference quotients, which can be used as approximations for the derivative of the function,

$$\Psi_j(t_{2i-1}) - \Psi_j(t_{2i}) = \frac{\Psi_j(t_{2i-1}) - \Psi_j(t_{2i-1}+h)}{h} h \approx -\Psi_j'(t_{2i-1})h. \tag{A.5}$$

The inner product can now be written as

$$\mathbf{v}_j^{\mathrm{T}}\mathbf{x} \approx -\frac{1}{\sqrt{N}}\sqrt{h}\sum_{i=1}^{N/2}\Psi_j'(t_{2i-1})h. \tag{A.6}$$

The terms of the sum are the values of the derivative of the wave function at positions separated by a distance $2h$. The sum can thus be related to the left Riemann sum of the function $\Psi_j'(x)$, which is an approximation for the integral of the function. Therefore, the sum can be approximated as

$$\sum_{i=1}^{N/2}\Psi_j'(t_{2i-1})2h \approx \int_{t_1}^{t_{N+1}}\Psi_j'(x)\,dx = \Psi_j(t_{N+1}) - \Psi_j(t_1). \tag{A.7}$$

The error terms arising from the approximations of the derivative with a difference quotient and the integral with a Riemann sum are now considered. The error of the left Riemann sum is estimated using the better known error bound of the trapezoidal rule, which is [42]

$$\epsilon_{trapez} = -\frac{b-a}{12}h^2 \max_{\xi\in[a,b]} f''(\xi), \tag{A.8}$$

where a and b are the start and end points of the integration interval. Since the trapezoidal rule for the integration is

$$I_{trapez} = h[\frac{1}{2}f(a) + f(x_1) + f(x_2) + \cdots + f(x_{n-1}) + \frac{1}{2}f(b)] \tag{A.9}$$

and the left Riemann sum is

$$I_{lsum} = h[f(a) + f(x_1) + f(x_2) + \cdots + f(x_{n-1})], \tag{A.10}$$

the absolute difference between the methods is

$$|I_{lsum} - I_{trapez}| = \frac{1}{2}h|f(a) - f(b)|. \tag{A.11}$$

Therefore, the absolute error of the left Riemann sum must be bounded by the error bound of the trapezoid rule and the absolute difference between the methods, i.e.

$$|\epsilon_{lsum}| \leq \frac{b-a}{12}h^2 \max_{\xi\in[a,b]} |f''(\xi)| + \frac{1}{2}h|f(a) - f(b)|. \tag{A.12}$$

The exact relation between the difference quotient and the derivative of the function can be derived using the Taylor's theorem and it results in

$$\frac{\Psi_j(t_{2i-1}+h) - \Psi_j(t_{2i-1})}{h} = \Psi_j'(t_{2i-1}) + \frac{1}{2}\Psi_j''(t_{2i-1})h + \frac{R_3(t_{2i-1})}{h}, \tag{A.13}$$

where $R_3(t_{2i-1})$ is the remainder of the Taylor expansion and can be written in Lagrange form as [43]

$$R_3(t_{2i-1}) = \frac{1}{3!}\Psi_j'''(\xi)h^3, \tag{A.14}$$

where $\xi \in [t_{2i-1}, t_{2i-1}+h]$. In Appendix B, it is proved that the Taylor series of the wave function converges for all h. Therefore, the remainder decreases as more terms are added to the expansion and adding more terms would improve the approximation. Adding the second order term and the remainder to Eq. (A.6) gives

$$\mathbf{v}_j^{\mathrm{T}}\mathbf{x} = -\frac{1}{\sqrt{N}}\sqrt{h}\left[\sum_{i=1}^{N/2}\Psi_j'(t_{2i-1})h + \frac{1}{2}\sum_{i=1}^{N/2}\Psi_j''(t_{2i-1})h^2 + \sum_{i=1}^{N/2}R_3(t_{2i-1})\right].$$

The sum of the second order derivatives is also a Riemann sum and can be approximated as an integral. Approximating the sums with integrals results in

$$\begin{aligned}\mathbf{v}_j^{\mathrm{T}}\mathbf{x} = -\frac{1}{\sqrt{N}}\sqrt{h}\Bigg[&\frac{1}{2}\left(\int_{t_1}^{t_{N+1}}\Psi_j'(x)\,dx + \epsilon_{lsum1}\right) + \\ &+ \frac{1}{4}h\left(\int_{t_1}^{t_{N+1}}\Psi_j''(x)\,dx + \epsilon_{lsum2}\right) + \sum_{i=1}^{N/2}R_3(t_{2i-1})\Bigg],\end{aligned}$$

where ϵ_{lsum1} and ϵ_{lsum2} are the errors of the first and the second Riemann sums respectively. The absolute value of the inner product is bounded by

$$\begin{aligned}|\mathbf{v}_j^{\mathrm{T}}\mathbf{x}| \leq\ & \frac{1}{\sqrt{N}}\sqrt{h}\Bigg[\frac{1}{2}|\Psi_j(t_{N+1}) - \Psi_j(t_1)| + \frac{1}{4}h|\Psi_j'(t_{N+1}) - \Psi_j'(t_1)| + \frac{1}{2}|\epsilon_{lsum1}| + \\ &+ \frac{1}{4}h|\epsilon_{lsum2}| + \sum_{i=1}^{N/2}|R_3(t_{2i-1})|\Bigg] \\ \leq\ & \frac{1}{\sqrt{N}}\Bigg[h^{1/2}\frac{1}{2}|\Psi_j(t_{N+1}) - \Psi_j(t_1)| + h^{3/2}\frac{3}{4}|\Psi_j'(t_{N+1}) - \Psi_j'(t_1)| + \\ &+ h^{5/2}\left(\frac{L}{4}\max_{\xi\in[t_1,t_{N+1}]}|\Psi_j'''(\xi)| + \frac{1}{4}|\Psi_j''(t_{N+1}) - \Psi_j''(t_1)|\right) + \\ &+ h^{7/2}\frac{L}{12}\max_{\xi\in[t_1,t_{N+1}]}|\Psi_j^{(4)}(\xi)|\Bigg],\end{aligned} \tag{A.15}$$

where the absolute errors of the left Riemann sums have been obtained from Eq. (A.12) and the sum of the remainders has been approximated by

$$\sum_{i=1}^{N/2} |R_3(t_{2i-1})| \leq \frac{1}{3!} h^3 \sum_{i=1}^{N/2} \max_{\xi \in [t_1, t_{N+1}]} |\Psi_j'''(\xi)| = \frac{L}{12} h^2 \max_{\xi \in [t_1, t_{N+1}]} |\Psi_j'''(\xi)|.$$

Appendix B

Convergence of wave function Taylor series

As it is not at all obvious that the Taylor expansion of the wave function which is used in Eq. (A.13) converges, a proof will be given here. The convergence of the expansion depends on the convergence of the Taylor series of the Lagrange functions, because the wave function consists of a finite sum of Lagrange functions. Therefore, it is enough to study the Taylor series of a single Lagrange function. A scaled Lagrange sinc function in 1-D can be written as

$$L_a(x) = \frac{1}{\sqrt{h}} \frac{\sin[\frac{\pi}{h}(x - x_a)]}{\frac{\pi}{h}(x - x_a)} = \frac{\sqrt{h}}{\pi} f(x)g(x), \tag{B.1}$$

where the Lagrange function has been separated into a product of two functions, $f(x) = \sin[\frac{\pi}{h}(x - x_a)]$ and $g(x) = (x - x_a)^{-1}$. Now, a Taylor expansion of the two functions is made around some mesh point x_b. The Taylor series of $f(x)$ is

$$f(x) = (-1)^{(x_b - x_a)/h} \left[\frac{\pi}{h}(x - x_b) - \frac{1}{3!} \left(\frac{\pi}{h}\right)^3 (x - x_b)^3 + O[(x - x_b)^5] \right], \tag{B.2}$$

which can be cast into the basic form of the sine function power series by changing the variable to $z = \frac{\pi}{h}(x - x_b)$. Since the power series of $\sin(z)$ is known to converge for all z, the Taylor series of $f(x)$ converges for all x. If the Taylor expansion is made around x_a, the expansion of the product $f(x)g(x)$ can be simply written as

$$\begin{aligned} f(x)g(x) &= \sum_{i=0}^{\infty} \frac{1}{(2i+1)!} \left(\frac{\pi}{h}\right)^{2i+1} (x - x_a)^{2i} \\ &= \left(\frac{\pi}{h}\right) \sum_{i=0}^{\infty} \frac{1}{2i+1} \frac{1}{(2i)!} \left(\frac{\pi}{h}\right)^{2i} (x - x_a)^{2i} \end{aligned} . \tag{B.3}$$

The expansion is almost the same as the power series of cosine, which converges for all x, except that each term is divided by $2i+1$. As $2i+1 \geq 1$, each term is smaller than in the cosine power series and, therefore, the Taylor expansion of $f(x)g(x)$ also converges for all x.

If the point around which the expansion is made is chosen so that $x_b \neq x_a$, the Taylor expansion of the product $f(x)g(x)$ is harder to obtain without calculating the derivatives one by one. In this case, the following theorem from real analysis can be used [44]. If $f(x) = \sum_{n=0}^{\infty} a_n x^n$ and $g(x) = \sum_{n=0}^{\infty} b_n x^n$ for $|x| < R$, then $f(x)g(x) = \sum_{n=0}^{\infty} c_n x^n$ for $|x| < R$, where

$$c_n = a_n b_0 + a_{n-1} b_1 + \cdots + a_0 b_n.$$

In other words, if the power series of $f(x)$ and $g(x)$ converge for $|x| < R$, also the power series of the product $f(x)g(x)$ converges for $|x| < R$. Since the power series of $f(x)$ converges for all x, the radius of convergence of the power series of $f(x)g(x)$ depends on that of the power series of $g(x)$. The Taylor series of $g(x)$ can be written as

$$g(x) = \frac{1}{x_b - x_a} \sum_{i=0}^{\infty} (-1)^i \left(\frac{x - x_b}{x_b - x_a} \right)^i . \tag{B.4}$$

As this is a geometric series, it converges if

$$\left| \frac{x - x_b}{x_b - x_a} \right| < 1 \quad \Rightarrow \quad |x - x_b| < |x_b - x_a|. \tag{B.5}$$

$|x_b - x_a|$ is a distance between two distinct mesh points and thus its minimum value is the mesh spacing h. Consequently, the minimum radius of convergence for the power series of $g(x)$ is $R = h$. According to the convergence theorem above, the power series of $f(x)g(x)$ thus converges for $|x - x_b| < h$. In the Taylor series in Eq. (A.13), $|x - x_b| = h$, however, and thus the convergence of the series is still uncertain. This limiting case where $|x - x_b| = |x_b - x_a|$ must be checked separately.

To see what happens when $|x - x_b| = |x_b - x_a|$, the terms in the power series of $f(x)g(x)$ must be calculated explicitely by multiplying the series expansions of $f(x)$ and $g(x)$. The product of these two series can be written as

$$f(x)g(x) = \left(\sum_{i=0}^{\infty} a_{2i+1} (x - x_b)^{2i+1} \right) \left(\sum_{j=0}^{\infty} b_j (x - x_b)^j \right) = \sum_{k=1}^{\infty} c_k (x - x_b)^k, \tag{B.6}$$

where

$$a_{2i+1} = (-1)^{(x_b - x_a)/h} \frac{(-1)^i}{(2i+1)!} \left(\frac{\pi}{h} \right)^{2i+1}, \quad b_j = (-1)^j (x_b - x_a)^{-j-1}.$$

c_k is the sum of all possible products of a_{2i+1} and b_j so that $2i + 1 + j = k$, i.e.

$$c_k = a_1 b_{k-1} + a_3 b_{k-3} + \cdots + \begin{cases} a_k b_0, & \text{if k is odd} \\ a_{k-1} b_1, & \text{if k is even} \end{cases} . \tag{B.7}$$

Using the definitions of a_{2i+1} and b_j above, the coefficients c_k can be written as

$$\begin{aligned} c_k &= (-1)^{(x_b - x_a)/h} \sum_{i=1}^{\lceil \frac{k}{2} \rceil} \frac{(-1)^{i-1}}{(2i-1)!} \left(\frac{\pi}{h} \right)^{2i-1} (-1)^{k-2i+1} \frac{1}{(x_b - x_a)^{k-2i+2}} \\ &= (-1)^{(x_b - x_a)/h} \frac{(-1)^k}{(x_b - x_a)^{k+1}} \sum_{i=1}^{\lceil \frac{k}{2} \rceil} \frac{(-1)^i}{(2i-1)!} \left(\frac{\pi}{h} \right)^{2i-1} (x_b - x_a)^{2i-1}. \end{aligned} \tag{B.8}$$

Now, the explicit form of the power series of $f(x)g(x)$ is

$$f(x)g(x) = (-1)^{(x_b-x_a)/h}\sum_{k=1}^{\infty}(-1)^k\frac{(x-x_b)^k}{(x_b-x_a)^{k+1}}\sum_{i=1}^{\lceil\frac{k}{2}\rceil}\frac{(-1)^i}{(2i-1)!}\left(\frac{\pi}{h}\right)^{2i-1}(x_b-x_a)^{2i-1}. \tag{B.9}$$

There are two possible values for x that satisfy $|x-x_b| = |x_b-x_a|$, $x = 2x_b - x_a$ and $x = x_a$. The convergence of the power series of $f(x)g(x)$ is first checked at $x = 2x_b - x_a$ using the explicit form of the series

$$\begin{aligned} &f(2x_b-x_a)g(2x_b-x_a) \stackrel{?}{=} \\ &\quad (-1)^{(x_b-x_a)/h}\frac{1}{x_b-x_a}\sum_{k=1}^{\infty}(-1)^k\sum_{i=1}^{\lceil\frac{k}{2}\rceil}\frac{(-1)^i}{(2i-1)!}\left(\frac{\pi}{h}\right)^{2i-1}(x_b-x_a)^{2i-1}. \end{aligned} \tag{B.10}$$

Due to the ceiling $\lceil\frac{k}{2}\rceil$, the inner sum is the same for each pair of subsequent odd and even values of k, i.e. $k = 2j-1$ and $k = 2j$. Since the terms in the outer sum have an alternating sign, the odd and even terms cancel each other out. Consequently, the series converges to zero if the inner sum converges to zero in the limit $k \to \infty$. Since the inner sum is clearly a partial sum of the series expantion of sine function, in the limit $k \to \infty$ it can be written as

$$\sum_{i=1}^{\infty}\frac{(-1)^i}{(2i-1)!}\left(\frac{\pi}{h}\right)^{2i-1}(x_b-x_a)^{2i-1} = -\sin[\frac{\pi}{h}(x_b-x_a)] = 0, \tag{B.11}$$

where $\sin[\frac{\pi}{h}(x_b-x_a)]$ is zero because $(x_b-x_a)/h$ is an integer. It can be concluded that the equivalence between $f(2x_b-x_a)g(2x_b-x_a)$ and the series expansion is justified because, assuming that $x_b \neq x_a$,

$$f(2x_b-x_a)g(2x_b-x_a) = \frac{\sin[\frac{\pi}{h}(2x_b-2x_a)]}{(2x_b-2x_a)} = 0,$$

which is the same value where the series expansion converges.

For the other possible value of x, $x = x_a$, the power series is

$$f(x_a)g(x_a) \stackrel{?}{=} (-1)^{(x_b-x_a)/h}\frac{1}{x_b-x_a}\sum_{k=1}^{\infty}\sum_{i=1}^{\lceil\frac{k}{2}\rceil}\frac{(-1)^i}{(2i-1)!}\left(\frac{\pi}{h}\right)^{2i-1}(x_b-x_a)^{2i-1}, \tag{B.12}$$

which is exactly same as that for $x = 2x_b - x_a$ except for the missing alternating sign in the outer sum. Therefore, none of the terms cancel now. The outer sum adds up inner sums where the index i runs up to $\lceil k/2\rceil$. Since the index k increased after each outer sum term, the inner sum runs further for each subsequent term of the outer sum. Consequently, the first term of the inner sum is added up in all M terms of the outer sum, the second term in $M-2$ terms of the outer sum, and so

on. The two sums can thus be simplified to

$$
\begin{aligned}
& f(x_a)g(x_a) \\
& \stackrel{?}{=} (-1)^{(x_b-x_a)/h} \quad \frac{1}{x_b-x_a} \lim_{M\to\infty} \left[\sum_{i=0}^{M} (M-2i) \frac{(-1)^{i+1}}{(2i+1)!} \left(\frac{\pi}{h}\right)^{2i+1} (x_b-x_a)^{2i+1} \right] \\
& = (-1)^{(x_b-x_a)/h} \quad \frac{1}{x_b-x_a} \left[\lim_{M\to\infty} (M) \sum_{i=0}^{\infty} \frac{(-1)^{i+1}}{(2i+1)!} \left(\frac{\pi}{h}\right)^{2i+1} (x_b-x_a)^{2i+1} - \right. \\
& \qquad\qquad \left. \sum_{i=0}^{\infty} 2i \frac{(-1)^{i+1}}{(2i+1)!} \left(\frac{\pi}{h}\right)^{2i+1} (x_b-x_a)^{2i+1} \right] \\
& = (-1)^{(x_b-x_a)/h} \quad \frac{1}{x_b-x_a} \left[\lim_{M\to\infty} (M)(-1) \sin\left[\frac{\pi}{h}(x_b-x_a)\right] - \right. \\
& \qquad\qquad \left. \sum_{i=0}^{\infty} (2i+1-1) \frac{(-1)^{i+1}}{(2i+1)!} \left(\frac{\pi}{h}\right)^{2i+1} (x_b-x_a)^{2i+1} \right] \\
& = (-1)^{(x_b-x_a)/h} \quad \left[\frac{\pi}{h} \sum_{i=0}^{\infty} \frac{(-1)^{i}}{(2i)!} \left(\frac{\pi}{h}\right)^{2i} (x_b-x_a)^{2i} + \right. \\
& \qquad\qquad \left. \frac{1}{x_b-x_a} \sum_{i=0}^{\infty} \frac{(-1)^{i+1}}{(2i+1)!} \left(\frac{\pi}{h}\right)^{2i+1} (x_b-x_a)^{2i+1} \right] \\
& = (-1)^{(x_b-x_a)/h} \quad \left[\frac{\pi}{h} \cos\left[\frac{\pi}{h}(x_b-x_a)\right] - \frac{1}{x_b-x_a} \sin\left[\frac{\pi}{h}(x_b-x_a)\right] \right] \\
& = \frac{\pi}{h}
\end{aligned}
\tag{B.13}
$$

where series expansions of $\sin\left[\frac{\pi}{h}(x_b-x_a)\right]$ and $\cos\left[\frac{\pi}{h}(x_b-x_a)\right]$ have been noticed and replaced with the function form. Since $(x_b-x_a)/h$ is an integer, $\cos\left[\frac{\pi}{h}(x_b-x_a)\right] = (-1)^{(x_b-x_a)/h}$. The series should converge to

$$
f(x_a)g(x_a) = \frac{\sin[\frac{\pi}{h}(x_a-x_a)]}{(x_a-x_a)} = \frac{\pi}{h}\,\mathrm{sinc}(0) = \frac{\pi}{h}, \tag{B.14}
$$

and thus the series is again equivalent to the product $f(x_a)g(x_a)$.

It has now been shown that the Taylor series of the Lagrange functions converge when $|x-x_b| \leq h$ regardless of the point at which the function is centered. x_b is the point around which the expansion is made. Consequently, the Taylor series of the wave function, which is a linear combination of the Lagrange functions that are centered at different mesh points, also converges when $|x-x_b| \leq h$. In the Taylor expansion of Eq. (A.13), $x-x_b$ is exactly h and, therefore, the series always converges, regardless of the mesh spacing h.

The fact that the Taylor expansion of the wave function around any mesh point x_b converges when $|x-x_b| \leq h$ is a curious one. It means that the radius of convergence depends on the mesh spacing h. The reason for this is that the convergence of the wave function Taylor series depends on that of the Lagrange function Taylor series.

Since the Lagrange functions are scaled with h, it is their natural length scale and the radius of convergence thus depends on it.

The length scale of the wave function depends only on physical variables. In this case, it depends on the confinement strength of the potential. As the convergence analysis shows, the calculated wave functions $\psi_j(x)$ seem to converge to the real ones as $h \to 0$. Therefore, they should be independent of the mesh spacing h.

Bibliography

[1] R. P. Feynman, "Simulating physics with computers," *International Journal of Theoretical Physics*, vol. 21, no. 6, pp. 467–488, 1982.

[2] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*. Cambridge University Press, 10th ed., 2010.

[3] P. W. Shor, "Algorithms for quantum computation: discrete logarithms and factoring," in *Proceedings of the 35th Annual Symposium on Foundations of Computer Science*, pp. 124–134, IEEE, 1994.

[4] L. K. Grover, "A fast quantum mechanical algorithm for database search," in *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing*, pp. 212–219, ACM, 1996.

[5] S. M. Reimann and M. Manninen, "Electronic structure of quantum dots," *Reviews of Modern Physics*, vol. 74, pp. 1283–1342, Nov 2002.

[6] R. C. Ashoori, "Electrons in artificial atoms," *Nature*, vol. 379, no. 6564, pp. 413–419, 1996.

[7] H. Ibach and H. Lüth, *Solid-State Physics: An Introduction to Principles of Materials Science*. Springer, 4th ed., 2009.

[8] A. Kumar, S. E. Laux, and F. Stern, "Electron states in a GaAs quantum dot in a magnetic field," *Physical Review B*, vol. 42, no. 8, p. 5166, 1990.

[9] T. Ando, A. B. Fowler, and F. Stern, "Electronic properties of two-dimensional systems," *Reviews of Modern Physics*, vol. 54, pp. 437–672, Apr 1982.

[10] T. D. Ladd, F. Jelezko, R. Laflamme, Y. Nakamura, C. Monroe, and J. L. O'Brien, "Quantum computers," *Nature*, vol. 464, no. 7285, pp. 45–53, 2010.

[11] J. Petta, A. Johnson, J. Taylor, E. Laird, A. Yacoby, M. Lukin, C. Marcus, M. Hanson, and A. Gossard, "Coherent manipulation of coupled electron spins in semiconductor quantum dots," *Science*, vol. 309, no. 5744, pp. 2180–2184, 2005.

[12] E. Nielsen, R. W. Young, R. P. Muller, and M. S. Carroll, "Implications of simultaneous requirements for low-noise exchange gates in double quantum dots," *Physical Review B*, vol. 82, p. 075319, Aug 2010.

[13] D. Loss and D. P. DiVincenzo, "Quantum computation with quantum dots," *Physical Review A*, vol. 57, no. 1, p. 120, 1998.

[14] A. M. Steane, "Overhead and noise threshold of fault-tolerant quantum error correction," *Physical Review A*, vol. 68, no. 4, p. 042322, 2003.

[15] D. Bruss and G. Leuchs, *Lectures on Quantum Information*. Wiley-Vch, 2007.

[16] A. V. Khaetskii, D. Loss, and L. Glazman, "Electron spin decoherence in quantum dots due to interaction with nuclei," *Physical Review Letters*, vol. 88, no. 18, p. 186802, 2002.

[17] J. Levy, "Universal quantum computation with spin-1/2 pairs and heisenberg exchange," *Physical Review Letters*, vol. 89, p. 147902, Sep 2002.

[18] J. Taylor, H.-A. Engel, W. Dür, A. Yacoby, C. Marcus, P. Zoller, and M. Lukin, "Fault-tolerant architecture for quantum computation using electrically controlled semiconductor spins," *Nature Physics*, vol. 1, no. 3, pp. 177–183, 2005.

[19] H. Bluhm, S. Foletti, I. Neder, M. Rudner, D. Mahalu, V. Umansky, and A. Yacoby, "Dephasing time of GaAs electron-spin qubits coupled to a nuclear bath exceeding 200 μs," *Nature Physics*, vol. 7, no. 2, pp. 109–113, 2010.

[20] R. L. Liboff, *Introductory Quantum Mechanics*. Addison-Wesley, 2003.

[21] F. Schwabl, *Advanced Quantum Mechanics*. Springer-Verlag, 2005.

[22] S. Foletti, H. Bluhm, D. Mahalu, V. Umansky, and A. Yacoby, "Universal quantum control of two-electron spin quantum bits using dynamic nuclear polarization," *Nature Physics*, vol. 5, no. 12, pp. 903–908, 2009.

[23] D. Baye, "Lagrange-mesh method for quantum-mechanical problems," *physica status solidi (b)*, vol. 243, no. 5, pp. 1095–1109, 2006.

[24] C. Schwartz, "High-accuracy approximation techniques for analytic functions," *Journal of Mathematical Physics*, vol. 26, no. 3, pp. 411–415, 1985.

[25] O. Kupiainen, "A Numerical Study of Double Quantum Dots Using the Lagrange Mesh Method," Master's thesis, University of Helsinki, 2010.

[26] C. Lanczos, "An iteration method for the solution of the eigenvalue problem of linear differential and integral operators," *Journal of Research of the National Bureau of Standards*, vol. 45, no. 4, 1950.

[27] G. H. Golub and C. F. Van Loan, *Matrix Computations*. The Johns Hopkins University Press, 2nd ed., Oct. 1996.

[28] C. Moler and C. Van Loan, "Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later," *SIAM Review*, vol. 45, no. 1, pp. 3–49, 2003.

[29] Y. Saad, "Analysis of some krylov subspace approximations to the matrix exponential operator," *SIAM Journal on Numerical Analysis*, vol. 29, no. 1, pp. 209–228, 1992.

[30] M. Hochbruck and C. Lubich, "On krylov subspace approximations to the matrix exponential operator," *SIAM Journal on Numerical Analysis*, vol. 34, no. 5, pp. 1911–1925, 1997.

[31] J. H. Wilkinson, ed., *The Algebraic Eigenvalue Problem*. New York, NY, USA: Oxford University Press, Inc., 1988.

[32] H. Karner, J. Schneid, and C. W. Ueberhuber, "Spectral decomposition of real circulant matrices," *Linear Algebra and its Applications*, vol. 367, pp. 301–311, 2003.

[33] J. C. R. Claeyssen and L. A. dos Santos Leal, "Diagonalization and spectral decomposition of factor block circulant matrices," *Linear Algebra and its Applications*, vol. 99, pp. 41 – 61, 1988.

[34] J. Tsai, L. Shieh, and R. Yates, "Fast and stable algorithms for computing the principal nth root of a complex matrix and the matrix sector function," *Computers & Mathematics with Applications*, vol. 15, no. 11, pp. 903–913, 1988.

[35] T. Hiltunen, J. Ritala, T. Siro, and A. Harju, "Non-adiabatic charge state transitions in singlet–triplet qubits," *New Journal of Physics*, vol. 15, no. 10, p. 103015, 2013.

[36] R. Hanson and G. Burkard, "Universal set of quantum gates for double-dot spin qubits with fixed interdot coupling," *Physical Review Letters*, vol. 98, p. 050502, Jan 2007.

[37] W. Coish and D. Loss, "Singlet-triplet decoherence due to nuclear spins in a double quantum dot," *Physical Review B*, vol. 72, no. 12, p. 125337, 2005.

[38] G. Ramon, "Electrically controlled quantum gates for two-spin qubits in two double quantum dots," *Physical Review B*, vol. 84, no. 15, p. 155329, 2011.

[39] D. Baye and P. H. Heenen, "Generalised meshes for quantum mechanical problems," *Journal of Physics A: Mathematical and General*, vol. 19, no. 11, p. 2041, 1986.

[40] D. Baye, M. Hesse, and M. Vincke, "The unexplained accuracy of the lagrange-mesh method," *Physical Review E*, vol. 65, p. 026701, Jan 2002.

[41] J. Särkkä and A. Harju, "Spin dynamics in a double quantum dot: Exact diagonalization study," *Physical Review B*, vol. 77, p. 245315, Jun 2008.

[42] E. Kreyszig, *Advanced Engineering Mathematics*. John Wiley, 9th ed., 2006.

[43] R. Courant and F. John, *Introduction to Calculus and Analysis*, vol. 1 of *Introduction to Calculus and Analysis*. Interscience Publishers, 1965.

[44] A. W. Knapp, *Basic Real Analysis*. Birkhäuser, 2005.